COWLES COMMISSION DISCUSSION PAPER: STATISTICS NO. 347

Combining Cross Section Data and Time Series

by Clifford Hildreth[1]

May 16, 1950

I

Two sorts of contributions to the problem of estimating economic relations from empirical data may be expected from the joint use of cross-section data and time series. The investigator can expect to work with large numbers of observations thus reducing his sampling errors and making tests of significance more powerful. He can also choose from a wider selection of statistical models thus having a better chance to construct a model that is both realistic and manageable. While only a few of the numerous problems of joint use of these types of data are considered in the present paper, both of these contributions are illustrated.

For purposes of comparison with other models to be discussed, the typical simultaneous equations model will be written -

$$(1.1)\quad \beta\, y^t + \Gamma z^t + \gamma^* = u^t .$$

β is a non-singular matrix of constant coefficients of current endogenous variables. y^t is a vector of the values taken by the observed current endogenous variables in the t-th observation. Γ is a matrix of constant coefficients of predetermined variables; z^t is a vector whose elements are values taken by the observed predetermined variables in the t-th observation. γ^* is a vector of the constant terms in the equations

---

1. I am indebted to Gurland, Hurwicz, Slater, and Sverdrup for helpful discussions of some of the problems encountered.

of the model. It could be included in Γ by making its elements the coefficients of a predetermined variable that takes the value 1 in each observation. However, for some of the comparisons to follow, it will be convenient to have these constant terms indicated separately. u^t is a vector of the values taken by the unobserved random disturbances in the equations in the t-th observation. For successive values of t, the u^t are assumed to represent successive independent drawings from a stable multivariate distribution. In the present discussion, disturbances are assumed to be normally distributed.

Most of the empirical applications thus far have used time series data, successive observations have represented values taken by the observed variables in successive time intervals. There have been some cross-section studies in which the same general model has been employed but in which different observations represented sets of values of the observed variables for different economic units (say households or firms) in the same time interval. The situation to be considered below is that in which the investigator has observations on each of a group of economic units for a number of time periods.

In this situation a model with the same statistical properties as (1.1) could be employed. This could be written -

$$(1.2)\quad B y^{it} + \Gamma z^{it} + \gamma^{**} = u^{it}.$$

y^{it} is a vector of the values taken by the current endogenous variables of the i-th economic unit and the t-th time period. Similarly z^{it} represents values of predetermined variables of the i-th unit and t-th time period. If it is also assumed that the u^{it} are drawings from a stable probability distribution with zero means and that a change in either superscript represents an independent drawing (i.e. u^{is} independent of u^{jt}, u^{is}, u^{js}), then model (1.2) has the same statistical properties as (1.1) and can be handled in the same fashion. In such a model it could happen that certain variables, say prices

of standardized, nationally marketed commodities, would be the same for all economic units at a given time and vary only over time. Such a variable will be denoted by y^{*t} or z^{*t}, possibly with a subscript to denote its position in the y^{it} or z^{it} vector. Similarly there might be variables such as formal education of an entrepreneur or head of household that would be constant over time but would vary over individual economic units. Such variables will be denoted by y^{i*} or z^{i*}. Some of the implications of these types of variables in models were discussed by the writer in Cowles Commission Discussion Paper: Statistics No. 333.

A model of the type of (1.2) might frequently be considered unrealistic for the following reason. It may be believed that there are unobserved individual characteristics which cause individuals to act differently and which are persistent over time. There may be other unobserved influences that affect individuals in pretty much the same way but change over time. Neither type of unobserved influence is expressed in (1.2). They could be included in a model of the following type -

$$(1.3)\quad \beta_y^{it} + \Gamma_z^{it} + \gamma^{i*} + \gamma^{*t} + \gamma^{**} = u^{it}.$$

γ^{i*} is a vector of constants that apply to the i-th individual and do not vary over time. Elements of γ^{*t} are constants characteristic of the t-th time period and do not vary from one individual to another. Other symbols are interpreted as before.

In (1.3), the γ^{i*} and γ^{*t} are interpreted as fixed variables; γ^{i*} having I different values and γ^{*t} having T different values in a sample of size IT containing observations on I individuals (economic units) in T time periods. I find it difficult to choose between the alternatives of allowing for these variations peculiar to individuals and variations peculiar to time through fixed parameters as in (1.3) or through random parameters. The latter choice could be expressed in a model of the form -

$$(1.4)\quad \beta_y^{it} + \Gamma_z^{it} + \gamma^{**} = u^{i*} + u^{*t} + u^{it}.$$

u^{i*} is regarded as a drawing from a normal multivariate population with zero means and

constant variances. u^{i*} is part of the disturbance in all equations relating to the i-th individual for every time period. u^{*t} is part of the disturbance in the equation for each individual in the t-th time period. u^{it} has the same interpretation as before and is distributed independently of u^{i*} and u^{*t}.

A generalization of model types (1.3) and (1.4) is given by -

$$(1.5)\ \beta_y^{it} + \gamma_z^{it} + \delta^{i*} + \delta^{*t} + \delta^{**} = u^{i*} + u^{*t} + u^{it}.$$

Each of the other types is a special case of (1.5) so this formulation might be used to derive tests of the special specifications in the others.

To consider the statistical aspects of the types of models indicated above I have found it convenient to first consider single-equation models of each type containing only one endogenous variable. As yet, I have not progressed much beyond this stage. However I believe that the single-equation results are of some interest in themselves - in some cases we may use single-equation models and in all cases our reduced form equations are of this type - and I expect that they will be helpful in considering simultaneous equations models.

II

The single-equation form of (1.2) will be written -

$$(2.1)\ \ y^{it} = \pi' z^{it} + \eta^{**} + v^{it}.$$

y^{it} is a scalar, π' a row vector of constant coefficients; z^{it} is a column vector and η^{**} is a scalar constant. v^{it} is a scalar non-observed random variable distributed independently of z^{it}. (2.1) may be regarded as a complete single-equation model or as one equation in the reduced form of a system like (1.2). Parameters of (2.1) can be estimated by least-squares.

To gain some notion of how the variances of estimates of the coefficients behave

under various assumptions about the type of data available, let us first assume that π and z^{it} are scalar. Then consider the following situations -

(a) The investigator has observations on a single individual for T time periods.

(b) The investigator has observations on the totals of the variables for a group of I individuals for each of T time periods. These totals over individuals will be designated as follows -

$$(2.2) \quad y^{It} = \sum_{i=1}^{I} y^{it}, \; z^{It} = \sum_{i=1}^{I} z^{it}, \; v^{It} = \sum_{i=1}^{I} v^{it} \; .$$

(c) Observations are available on each of I individuals in each of T time periods.

The data in situations (a) and (b) are time series. In situations (c) we have a combination of cross-section data and time series. It is not necessary to discuss the pure cross-section case (I individuals, one time period) in this paper since variations over time and variations over individuals are treated symmetrically in all of the models presented.

In situation (a) we know that -

$$(2.3) \quad \hat{\pi}_a = \frac{\sum_{t=1}^{T} (y^{it} - \frac{1}{T} y^{iT}) (z^{it} - \frac{1}{T} z^{iT})}{\sum_{t=1}^{T} (z^{it} - \frac{1}{T} z^{iT})^2}$$

and the variance of this estimate is -

$$(2.4) \quad V(\hat{\pi}_a) = \frac{\sigma^2}{\sum_{t=1}^{T} (z^{it} - \frac{1}{T} z^{iT})^2}$$

where $z^{iT} = \sum_{t=1}^{T} z^{it}$ and $y^{iT} = \sum_{t=1}^{T} y^{it}$ and σ^2 is the variance of v^{it}

Similarly -

$$(2.5) \quad V(\hat{\pi}_b) = \frac{I\sigma^2}{\sum_{t=1}^{T} (z^{It} - \frac{1}{T} z^{IT})^2} \qquad \text{and}$$

$$(2.6)\quad V(\hat{\pi}_c) = \frac{\sigma^2}{\sum_{t=1}^{T}\sum_{i=1}^{I}\left(z^{it} - \frac{1}{IT} z^{IT}\right)^2}$$

are the variances of the least squares estimates of π in situations (b) and (c) respectively. Clearly the relations among the three variances depend on the nature of the variation in z^{it}. Consider the following expression of the variation in z^{it} -

$$(2.7)\quad z^{it} = s^{**} + s^{i*} + s^{*t} + s^{it}$$

$$s^{**} = \frac{z^{IT}}{IT}, \quad s^{i*} = \frac{z^{iT}}{T} - s^{**}, \quad s^{*t} = \frac{z^{It}}{I} - s^{**}.$$

Variances of the estimates can be written -

$$(2.8)\quad V(\hat{\pi}_a) = \frac{\sigma^2}{\sum_t (s^{*t} + s^{it})^2}$$

$$V(\hat{\pi}_b) = \frac{\sigma^2}{I \sum_t (s^{*t})^2}$$

$$V(\hat{\pi}_c) = \frac{\sigma^2}{T \sum_i (s^{i*})^2 + I \sum_t (s^{*t})^2 + \sum_i \sum_t (s^{it})^2}.$$

If $z^{it}=z^{*t}$, then $s^{i*}=s^{it}=0$ and $z^{*t}=s^{**}+s^{*t}$. In this case $V(\hat{\pi}_b)$ and $V(\hat{\pi}_c)$ are equal to each other and to $\frac{1}{I} V(\hat{\pi}_a)$. Intuitively it would seem that if variables z^{it} and z^{*t} both entered our equation, then the additional information on z^{it} in situation (c) should, in general, enable us to estimate coefficients of both z^{it} and z^{*t} more accurately than in situation (b). This is verified by considering the case -

$$(2.9)\quad y^{it} = \pi_1 z_1^{it} + \pi_2 z_2^{*t} + \eta^{**} + v^{it}.$$

Let $z_1^{it} = s_1^{**} + s_1^{i*} + s_1^{*t} + s^{it}$ and $z_2^{*t} = s_2^{**} + s_2^{*t}$.

We have the following variances of estimates of π_1 and π_2 in our three situations -

(2.10)

$$V(\hat{\pi}_{1a}) = \frac{\sigma^2 \sum_t (s_2^{*t})^2}{\sum_t (s_1^{*t} + s_1^{it})^2 \cdot \sum_t (s_2^{*t})^2 - \left\{\sum_t s_2^{*t} (s_1^{*t} + s_1^{it})\right\}^2}$$

$$V(\hat{\pi}_{2a}) = \frac{\sigma^2 \sum_t (s_1^{*t} + s_1^{it})^2}{\sum_t (s_1^{*t} + s_1^{it})^2 \cdot \sum_t (s_2^{*t})^2 - \left\{\sum_t s_2^{*t} (s_1^{*t} + s_1^{it})\right\}^2}$$

$$V(\hat{\pi}_{1b}) = \frac{\sigma^2 \sum_t (s_2^{*t})^2}{I \sum_t (s_1^{*t})^2 \cdot \sum_t (s_2^{*t})^2 - I \left(\sum_t s_1^{*t} s_2^{*t}\right)^2}$$

$$V(\hat{\pi}_{2b}) = \frac{\sigma^2 \sum_t (s_1^{*t})^2}{I \sum_t (s_1^{*t})^2 \cdot \sum_t (s_2^{*t})^2 - I \left(\sum_t s_1^{*t} s_2^{*t}\right)^2}$$

$$V(\hat{\pi}_{1c}) = \frac{\sigma^2 \sum_t (s_2^{*t})^2}{\left\{T \sum_i (s_1^{i*})^2 + I \sum_t (s_1^{*t})^2 + \sum_i \sum_t (s_1^{it})^2\right\} \sum_t (s_2^{*t})^2 - I \left(\sum_t s_1^{*t} s_2^{*t}\right)^2}$$

$$V(\hat{\pi}_{2c}) = \frac{\sigma^2 \left\{T \sum_i (s_1^{i*})^2 + I \sum_t (s_1^{*t})^2 + \sum_i \sum_t (s_1^{it})^2\right\}}{\left\{T \sum_i (s_1^{i*})^2 + I \sum_t (s_1^{*t})^2 + \sum_i \sum_t (s_1^{it})^2\right\} \sum_t (s_2^{*t})^2 - I \left(\sum_t s_1^{*t} s_2^{*t}\right)^2}$$

Comparison of $V(\hat{\pi}_{2b})$ and $V(\hat{\pi}_{2c})$ shows the latter to be smaller unless z_1^{it} and z_2^{*t} are orthogonal in which case they are equal.

III

Using notation already established, the single-equation form of (1.3) is

$$(3.1)\quad y^{it} = \pi' z^{it} + \eta^{i*} + \eta^{*t} + \eta^{**} + v^{it}.$$

This differs from (2.1) in the inclusion of the I + T additional parameters, η^{i*} and η^{*t}. To obtain estimates of parameters in this equation I have used the method of maximum likelihood. The log of the likelihood function (except for a constant term) for IT observations i=1...I, t=1...T is

$$(3.2)\quad \ell = -IT \log\sigma - \frac{1}{2\sigma^2} \sum_t \sum_i (y^{it} - \pi' z^{it} - \eta^{**} - \eta^{i*} - \eta^{*t})^2$$

Differentiating with respect to σ^2, π', η^{**}, η^{i*}, and η^{*t} leads to the estimates indicated below. $\sum_i \eta^{i*}$ and $\sum_t \eta^{*t}$ are arbitrarily assumed equal to zero to avoid lack of identification of η^{**}, η^{i*}, η^{*t}.

$$(3.3)\quad \hat{\sigma}^2 = \frac{1}{IT} \sum_t \sum_i (y^{it} - \hat{\pi}' z^{it} - \hat{\eta}^{**} - \hat{\eta}^{i*} - \hat{\eta}^{*t})^2$$

$$\hat{\eta}^{**} = \frac{1}{IT} (y^{IT} - \hat{\pi}' z^{IT})$$

$$\hat{\eta}^{i*} = \left(\frac{y^{iT}}{T} - \frac{y^{IT}}{IT}\right) - \hat{\pi}' \left(\frac{z^{iT}}{T} - \frac{z^{IT}}{IT}\right)$$

$$\hat{\eta}^{*t} = \left(\frac{y^{It}}{I} - \frac{y^{IT}}{IT}\right) - \hat{\pi}' \left(\frac{z^{It}}{I} - \frac{z^{IT}}{IT}\right)$$

$\hat{\Pi}' = \overset{o}{M}_{yz} \overset{o}{M}_{zz}^{-1}$ where

$$\overset{o}{M}_{yz} = (\overset{o}{m}_{yz_1} \quad \overset{o}{m}_{yz_2} \quad \cdots \quad \overset{o}{m}_{yz_n})$$

$$\overset{o}{M}_{zz} = \begin{pmatrix} \overset{o}{m}_{z_1z_1} & \overset{o}{m}_{z_1z_2} & \cdots & \overset{o}{m}_{z_1z_n} \\ \overset{o}{m}_{z_2z_1} & \overset{o}{m}_{z_2z_2} & \cdots & \overset{o}{m}_{z_2z_n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \overset{o}{m}_{z_nz_1} & \overset{o}{m}_{z_nz_2} & \cdots & \overset{o}{m}_{z_nz_n} \end{pmatrix} \quad \text{and}$$

$$\overset{o}{m}_{yz_1} = \sum_i \sum_t (y^{it} - \frac{y^{It}}{I} - \frac{y^{iT}}{T} + \frac{y^{IT}}{IT}) \quad (z^{it} - \frac{z^{It}}{I} - \frac{z^{iT}}{T} + \frac{z^{IT}}{IT})$$

and similarly for the other sample moments.

It seems quite possible to me that the covariance matrix of the estimates, $\hat{\Pi}'$, will be given by $\overset{o}{M}_{zz}^{-1} \sigma^2$ and also that the limited information method of estimation of a structural system like (1.3) would be similar to the usual case except that moments $\overset{o}{m}$ would be used instead of the usual sample moments. I have not had time to check either of these propositions.

It is worth noting that the usual time series model and the usual cross-section model are inconsistent with the model just discussed. If we regard (3.1) as explaining the generation of the observed y^{it}, then the corresponding time series equation in the aggregate variables is -

$$(3.4) \quad y^{It} = \Pi' z^{It} + \eta^{I*} + I\eta^{*t} + I\eta^{**} + v^{It}.$$

If $\tilde{\Pi}'$ represents the estimate of Π' obtained by ordinary least squares, we have the following -

(3.5) $\tilde{\Pi}' = M_{yz}^{(I)} M_{zz}^{(I)-1}$ where, for instance,

(3.6) $m_{yz_1}^{(I)} = \sum_t (y^{It} - \frac{y^{IT}}{T})(z^{It} - \frac{z^{IT}}{T})$.

To indicate the bias in these estimates we note that -

(3.7) $E(M_{yz}^{(I)}) = \Pi' M_{zz}^{(I)} + I \cdot M_{\eta z}^{(I)}$ where

$$m_{\eta z_1}^{(I)} = \sum_t \eta^{*t} (z_1^{It} - \frac{z^{IT}}{T}), \text{ etc.}$$

(3.8) $E(\tilde{\Pi}') = \Pi' + I \cdot M_{\eta z}^{(I)} M_{zz}^{(I)-1}$.

Since the estimates in (3.3) are maximum likelihood estimates, the investigator could readily form likelihood ratio tests of various hypotheses. If, for instance, he tested the deviations of the η^{*t} from zero and found the deviations significant, he would have evidence that the usual time series model would lead to biased estimation.

IV

(4.1) $y^{it} = \Pi' z^{it} + \eta^{**} + v^{i*} + v^{*t} + v^{it}$

is the single equation version of (1.4). v^{i*} is assumed to represent an independent drawing from a probability distribution with zero mean and finite variance for each individual in the sample. v^{i*} is stable over time. Analogously, v^{*t} is stable over individuals and varies randomly over time. As before, v^{it} is assumed to vary randomly over both individuals and time. v^{it} is assumed to be independent of v^{i*} and v^{*t}.

(4.1) differs from (2.1) in that it does provide for the possibilities that there may be persistent unobserved individual characteristics giving rise to differences in individual economic behavior and there may be unobserved properties of time intervals

giving rise to differences in behavior that are constant over individuals but changing over time. (4.1) differs from (3.1) in introducing these possibilities through random variables rather than fixed variables. Unlike (3.1), (4.1) is consistent with the usual treatment of purely cross-section data or pure time series.

On grounds of realism I find it difficult to establish a preference as between (4.1) and (3.1) so I would like to be able to handle either and to develop statistical tests of their special assumptions.

Maximum likelihood did not work at all well in this case. The estimating equations are difficult to derive and appear to be highly non-linear in the unknown parameters. The reason is that successive values taken by the disturbance are no longer independent drawings from a stable population.

It does happen, however, that the estimation equations derived in section III provide unbiased estimates of the coefficients in this model. This will be shown in the next section for a slightly more general model that includes (4.1) as a special case.

V

A single equation model corresponding to (1.5) could be written -

$$(5.1)\quad y^{it} = \pi' z^{it} + \eta^{i*} + \eta^{*t} + \eta^{**} + v^{i*} + v^{*t} + v^{it}.$$

An attempt to handle this case by maximum likelihood would involve the same difficulties as (4.1). To show that the estimation equations of section III provide unbiased estimates of the coefficients in (5.1), let us first write -

$$(5.2)\quad y^{it} = \pi' z^{it} + w^{i*} + w^{*t} + w^{it} \qquad \text{where}$$

$$w^{i*} = \eta^{i*} + v^{i*}$$

$$w^{*t} = \eta^{*t} + v^{*t}$$

$$w^{it} = \eta^{**} + v^{it}.$$

From (3.3) we have -

$$(5.3)\quad \hat{\pi}' = \mathfrak{M}_{yz}\,\mathfrak{M}_{zz}^{-1} \quad \text{where}$$

$$(5.4)\quad \mathfrak{M}_{yz} = \sum_i \sum_t \left(y^{it} - \frac{y^{It}}{I} - \frac{y^{iT}}{T} + \frac{y^{IT}}{IT}\right)\left(z^{it} - \frac{z^{It}}{I} - \frac{z^{iT}}{T} + \frac{z^{IT}}{IT}\right)'$$

$$= \sum_i \sum_t y^{it}(z^{it})' - \frac{1}{I}\sum_t y^{It}(z^{It})' - \frac{1}{T}\sum_i y^{iT}(z^{iT})' + \frac{1}{IT}y^{IT}(z^{IT})'.$$

Consider the four terms in the last expression separately, we have -

$$(5.5)\quad \sum_i \sum_t y^{it}(z^{it})' = \pi' \sum_i \sum_t z^{it}(z^{it})' + \sum_i w^{i*}(z^{iT})' + \sum_t w^{*t}(z^{It})' + \sum_i \sum_t w^{it}(z^{it})'$$

$$-\frac{1}{I}\sum_t y^{It}(z^{It})' = -\pi'\frac{1}{It}\sum z^{It}(z^{It})' - \frac{1}{I}w^{I*}(z^{IT})' - \sum_t w^{*t}(z^{It})' - \frac{1}{I}\sum_t w^{It}(z^{It})'$$

$$-\frac{1}{T}\sum_i y^{iT}(z^{iT})' = -\pi'\frac{1}{T}\sum_i z^{iT}(z^{iT})' - \sum_i w^{i*}(z^{iT})' - \frac{1}{T}w^{*T}z^{IT} - \frac{1}{T}\sum_i w^{iT}(z^{iT})'$$

$$\frac{1}{IT}y^{IT}(z^{IT})' = \pi'\frac{1}{IT}z^{IT}(z^{IT})' + \frac{1}{I}w^{I*}(z^{IT})' + \frac{1}{T}w^{*T}(z^{IT})' + \frac{1}{IT}w^{IT}(z^{IT})'$$

If we now add each column, we get from the first column to the right of the equality signs -

$$(5.6)\quad \pi'\left\{\sum_i \sum_t z^{it}(z^{it})' - \frac{1}{I}\sum_t z^{It}(z^{It})' - \frac{1}{T}\sum_i z^{iT}(z^{iT})' + \frac{1}{IT}z^{IT}(z^{IT})'\right\} = \pi'\mathfrak{M}_{zz}$$

From the second column to the right of the equality signs we get zero and from the third column we also get zero. From the last column we get -

$$(5.7)\quad \sum_i \sum_t w^{it}(z^{it})' - \frac{1}{I}\sum_t w^{It}(z^{It})' - \frac{1}{T}\sum_i w^{iT}(z^{iT})' + \frac{1}{IT}w^{IT}(z^{IT})'$$

$\mathfrak{M}_{yz}$ is the sum of (5.6) and (5.7). If we take expected values, (5.6) is unchanged. The expected value of (5.7) is

$$(5.8)\quad \eta^{**}(z^{IT})' - \eta^{**}(z^{IT})' - \eta^{**}(z^{IT})' + \eta^{**}(z^{IT})' = 0.$$

We thus have -

$$(5.9)\quad E\,(\overset{o}{M}_{yz}) = \pi' \overset{o}{M}_{zz} \quad \text{and}$$

$$(5.10)\quad E\,(\overset{A'}{\pi}) = \pi$$

It can also be shown that -

$$(5.11)\quad E\,(\hat{\eta}^{**}) = \eta^{**}$$

$$E\,(\hat{\eta}^{i*}) = \eta^{i*}$$

$$E\,(\hat{\eta}^{*t}) = \eta^{*t}$$

I have not yet attempted to derive tests of the special assumptions in the other models using (5.1) as . However it would seem that it should be possible to derive such tests.