Confluence Analysis When the Model Contains Both Shocks and Errors

by Olav Reiersøl

I. Introduction

1. Notation and terminology.

We shall introduce the following notations for sequences of numbers;

$$J = [1 \ 2 \ \dots \ n_J]$$
$$I = [1 \ 2 \ \dots \ n_I]$$
$$G = [1 \ 2 \ \dots \ n_g]$$

We shall suppose that $n_J \geqq n_I \geqq n_g$, so that $J \supset I \supset G$. If S is any sequence containing the sequence S', the difference $S-S'$ will be defined as the sequence which contains the elements of S which are not in S', the elements being taken in the same order as in S. If $S' = [s_1' \ s_2' \ \dots \ s_n']$ and $S'' = [s_1'' \ s_2'' \ \dots \ s_n'']$, the sequence $[s_1' \ \dots \ s_n' \ s_1'' \ \dots \ s_n'']$ will be denoted by $S'S''$.

An element of a sequence will be denoted by the corresponding small letter. For instance the subscript g shall be taken to mean an element in the sequence G. The number of elements in a sequence will be denoted by n with the small letter of the sequence as subscript. A letter with a capital letter as subscript will denote a vector whose elements have as subscripts the elements of the sequence, for instance

$$x_G = [x_1 \ x_2 \ \dots \ x_{n_g}]$$

A similar notation will be used for matrices, for instance

$$\alpha_{GI} = \begin{bmatrix} \alpha_{11} & \text{---------} & \alpha_{1n_i} \\ \text{-----------------------} \\ \alpha_{n_g 1} & \text{-------} & \alpha_{n_g n_i} \end{bmatrix}$$

Let x_J be a set of observable variables which are (or may be) subject to errors. Let the error-free part of x_J be y_J and let the error be

$$(1) \quad v_J = x_\tau - y_J$$

Let $f_G(y_t)$ denote a vector function of y_t with n_y components, and let $y_{t\cdot J}$ denote the vector y_t at time t. We shall suppose that $y_{t\cdot J}$ satisfies the following equations

$$(2) \quad f_G(y_{t\cdot I}, y_{t-1\,I}, \ldots, y_{t-n_u\,I}) = u_{t\,G}$$

and

$$(3) \quad f_{J-G}(y_{t\cdot J-G}, y_{t-1\cdot J-G}, \ldots, y_{t-1\,G}, y_{t-2\,G} \ldots) = u_{t\,J-G}$$

Equations (2) and (3) will be supposed to be such that they can be solved uniquely in terms of $y_{t\cdot G}$ and $y_{t\cdot J-G}$ respectively. u_G, u_{J-G} and v_J will be supposed to be random vectors with mean value zero. Before making any further assumptions about the nature of these random vectors we shall distinguish between two cases:

Case A: The variables y_G occur only at time t in equation (2) and do not occur at all in equation (3). In this case we assume that

(4) $u_{t\cdot G}$ is independent of $u_{t-u\,J-G}$ when $u = 0, 1, 2,$

Case B: At least one variable y_J occurs with a time lag either in equation (2) or in equation (3). In this case we shall still make assumption (4) and in addition to that also the assumption

(5) $u_{t\cdot G}$ is independent of $u_{t-u\,G}$ when $u = 1, 2, \cdots$

For both cases we have

(6) $u_{t\,G}$ is independent of all variables y in equation (2) except $y_{t\cdot G}$

We shall suppose that, at least for the time being, we are only interested in getting information about Equation (2). This vector equation will be called the <u>investigational system of equations</u> and the set of variables occurring in it will be called the <u>investigational set of variables</u>. System (3) will be called the <u>outside system</u> and the variables will be called <u>outside variables</u>. The variables will be called <u>post-determined variables</u>

and the other variables y in system (2) will be called pre-determined.
The same distinctions will be applied to the corresponding x's.

We shall now change our notations for the variables, listing each
variable $y_{t-u.I}$ as a new variable $y_{t.I}$. We shall use the notation I for the
sequence of integers from 1 to the total number of variables after this renumber-
ing, and the sequence I - G shall be denoted by K. We shall suppose that system
(2) is linear and shall write it in matrix form:

(7) $\quad \alpha_{G.I} \, y_I^t = u_G^t$

where $\alpha_{G.G}$ is supposed to be non-singular.

We shall suppose that the variables involved are stationary. This
is no necessary assumption but will be made for convenience. We shall make
no assumption about system (3) which is not already contained in this assump-
tion and assumption (4). We shall suppose that

$$E(y_I) = E(x_I) = 0$$

2. Reduction to a pure error model.

Let the reduced form of 7 be

(8) $\quad y_G^t + \pi_{G.K} \, y_K^t = \alpha_{G.G}^{-1} \, u_G^t$

We shall now introduce the structural part s_I and the disturbance
w_I of the vector x_I by the following definitions:

(9) $\quad s_G^t = y_G^t - \alpha_{G.G}^{-1} \, u_G^t$

(10) $\quad s_K = y_K$

and

(11) $\quad w_I = x_I - s_I$

It should be noted that if y_k is a lagged value of y_g, s_k will no
longer be a lagged value of s_g.

From (1), (9), (10) and (11) we obtain

(12) $\quad w_G^t = v_G^t + \alpha_{G.G}^{-1} \, u_G^t$

and

$$(13) \quad w_K^- = v_K^-$$

From (8) and (9) we obtain

$$(14) \quad z_G' + \pi_{G \cdot K} \, z_K' = 0$$

so that our model in a formal way is reduced to a pure error model. This kind of model has been studied by several authors. (Ref $[2] - [10]$) But not all the work done on pure error models is applicable here, because some of it presupposes that all the w's are uncorrelated and there is no reason to make that assumption here, except for the case when there is only one post-determined variable. The previous work on the pure error model has partly been concerned with models leading to identifiability[1] of at least some of the parameters, partly it has been concerned with finding bounds for the parameter point in cases where there is no identifiability. The latter problem has previously been studied only for the pure error model with all errors uncorrelated. The main purpose of this paper will be to find bounds for the parameter point in cases of non-identifiability. I shall, however, mention briefly some previous work on identifiable models.

3. Models where the coefficient matrix of the system is identifiable.

We shall particularly consider the case of one post-determined variable and one equation in the investigational system

$$(15) \quad \alpha_I \, z_I' = 0$$

Suppose that we have another set of variables x_Q which are correlated with the structural parts z_I but not with the disturbances w_I. (Ref $[6]$, last part of Section 11; Ref. $[4]$, Section 2; Ref $[7]$, Section 5.) We shall call the set x_Q an _instrumental set of variables_. From (11) and (15) we obtain

$$(16) \quad \alpha_I \, E(x_I' \, x_Q) = 0$$

The vector α_I will be identifiable if the rank of $E(x_I' \, x_Q)$ is n_1-1. The

---

[1]Identifiable will in this paper always be taken to mean what is called uniquely identifiable in Cowles Commission Monograph No. 10.

rank cannot be greater than n_1-1 if our model is correct. In Ref. [7] I have considered a model which in one way is more general, in another way more special. It is more general in the sense that I have not supposed all covariances $E(w_1 x_q)$ to be zero. It is more special in the sense that the number of instrumental variables is equal to the number of investigational variables. In this model we usually do not have identifiability but may be able to find bounds for the structural vector α_I. In the case where all covariances $E(w_1 x_q)$ are equal to zero, my theorems about bounds in Ref. [6] and Ref. [7] reduce to theorems about identifiability of the vector α_I provided that certain matrices are different from zero.

As instrumental variables we may introduce some of the variables x_{J-I} or functions of them. Since we have assumed u_G independent of u_{J-G}, u_G may be supposed to be independent of all such variables. If we also assume the error v_I to be independent of the error v_{J-I}, then w_I will be independent of x_{J-I}. If the set x_{J-I} contains lagged values of the variables x_I, this leads us to introduce lagged values of the investigational variables as instrumental variables. We may also use lagged values of the variables x_I not occurring in the set x_{J-I}. (Compare Ref [6], Section 4 and Ref. [7], Chapter VI).

Other models leading to identification have been studied by Wald (Ref. [10], Geary (Ref. [3]), and Tintner (Ref. [8] and [9]).

In spite of the different models leading to identifiability, we still have cases, important in practice, where there is no identifiability and we are now going to consider such cases.

II Models Where the Parameter Point Is Not Identifiable But Surroundable by a Bounded Region.

4. Definitions and assumptions.

We shall introduce notations for the following covariance matrices:

$$\mu_{I \cdot I} = E(\kappa_I' \kappa_I)$$

$$\nu_{I \cdot I} = E(z_I' z_I)$$

and

$$\lambda_{I \cdot I} = E(w_I' \, w_I)$$

We shall make the following assumptions

(17) $\quad E(v_I' \, y_I) = 0$

(18) $\quad E(v_I' \, u_G) = 0$

(19) $\quad E(v_G' \, v_K) = 0$

(20) $\quad E(v_K' \, v_K)$ is diagonal

(21) $\quad v_V = 0 \quad$ where $V \subset K$

We shall write $L = K - V$ and $H = I - V$. As a particular case V may be empty
so that $L = K$ and $H = I$.

From assumptions (6) and (17) follows that

(22) $\quad \mu_{I \cdot I} = v_{I \cdot I} + \lambda_{I \cdot I}$

From assumptions (18) - (21) follows that $\lambda_{I \cdot I}$ is of the form

$$(23) \quad \lambda_{I \cdot I} = \begin{bmatrix} \lambda_{G \cdot G} & 0 & 0 \\ 0 & \lambda_L & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where we have written λ_L instead of $\lambda_{L \cdot L}$ to point out that it is a
diagonal matrix.

At this point we shall also introduce some mathematical concepts which
will be used in the following:

A <u>polytope</u> is the extension to n dimensions of a polygon in 2 dimensions
and a polyhedron in 3 dimensions.

An <u>orthant</u> is the extension to n dimensions of a quadrant in two
dimensions and an octant in 3 dimensions.

A <u>p-flat</u> is a linear manifold of p dimensions.

A matrix $\mu_{G \cdot I}$ will be said to have <u>compatible signs</u> if there exist
diagonal matrices $\mathcal{E}_G$ and $\mathcal{E}_I$ such that $\mathcal{E}_G \mu_{G \cdot I} \mathcal{E}_I$ is positive (i.e. that
all elements of this matrix are positive.)

5. The case of one single post-determined variable.

We shall now write the single equation of the investigational system in the form

$$(24) \quad z_1 + \pi_K \; z_K' = 0$$

Let $\hat{\mu}_{II}$ be the adjoint of μ_{II}, and let β_{II} be the matrix obtained from $\hat{\mu}_{II}$ by dividing each row by the first element of that row. Let the vectors β_{hK} and the vector π_K be interpreted as points in an (n_1-1)-dimensional space. The convex polytope with the n_h points β_{hK} as vertices will be denoted by $B_{H \cdot K}$.

We have the following theorems:

Theorem I: _If the matrix_ $\beta_{H \cdot K}$ _has compatible signs, the structural point_ π_K _is confined to the polytope_ $B_{H \cdot K}$.

Theorem II: _If_ z_i _and_ w_i _are normally distributed and if_ $z_{tI} w_{tI}$ _is independent of_ $z_{t'I} w_{t'I}$ _when_ $t \neq t'$, _then each point in the polytope represents a possible structure under the given assumptions and the given probability distribution of_ κ_I.

If the assumption about normality and independence does not hold good, we can still state that the bounds given in Theorem I are the best possible if we have no other relevant information than that given by the covariance matrix μ_{II} and the conditions of the theorem.

All these statements follow immediately from Theorem 17.3 in Ref [7]. The special case of Theorem I when H = I was first given in a somewhat different form by Koopmans (Ref. [5], p. 101). Frisch had previously proved the theorem in the case of two variables and had anticipated the general theorem and used it as an argument for his bunch analysis. (Ref. [2], Section 9).

6. The general case.

We shall find bounds for the row-vectors of the reduced form matrix $\pi_{G \cdot K}$. We shall consider one particular equation

$$(25) \quad z_1 + \pi_{g \cdot K} \; z_K' = 0$$

If we consider the set of variables k_{gK} the matrix λ_{gK} is diagonal so that Theorem I may be applied. Let us write μ_{gK*} instead of $\mu_{gK \cdot gK}$ and let $\beta^{(gK)}_{gK*}$ be the matrix which is obtained from $\hat{\mu}_{gK*}$ after dividing each row by the first element of this row. If the matrix $\beta^{(gK)}_{gL*}$ has compatible signs, the structural point $\bar{k}_{\cdot K}$ is confined to the polytope $B^{(gK)}_{gL \cdot K}$. If we do not insist on the narrowest possible bounds we can therefore still apply Frisch's bunch analysis, but only to the sets containing all pre-determined variables and one single post-determined variable.

In the following we shall assume that

(26) there exists no admissible point λ_K such that $v_{K \cdot K}$ is singular. This assumption will certainly hold good if

(27) there exists at least one g such $\left(\hat{\mu}_{gK*}\right)_{gL*}$ has compatible signs. This follows from a generalization of Theorems 12 and 13 in Ref. $[6]$.
Let us now again consider the whole reduced form (14). Postmultiplying by Z_I and taking the mathematical expectation, we obtain

$$(28) \quad v_{G \cdot I} + \pi_{G \cdot K} \, v_{K \cdot I} = O$$

which may be split into two equations

$$(29) \quad v_{G \cdot G} + \pi_{G \cdot K} \, \mu_{K \cdot G}$$

and

$$(30) \quad \mu_{G \cdot K} + \pi_{G \cdot K} \, v_{K \cdot K} = O$$

From (29) and (30) we obtain

$$(31) \quad \pi_{G \cdot K} = -\mu_{G \cdot K} \, v_{K \cdot K}^{-1}$$

$$(32) \quad v_{G \cdot G} = \mu_{G \cdot K} \, v_{K \cdot K}^{-1} \, \mu_{K \cdot G}$$

and

$$(33) \quad \lambda_{G \cdot G} = \mu_{G \cdot G} - \mu_{G \cdot K} \, v_{K \cdot K}^{-1} \, \mu_{K \cdot G}$$

Using (33) we obtain the following identity

$$(34) \quad \begin{bmatrix} I_G & \mu_{G \cdot K} v_{K \cdot K}^{-1} \\ O & I_K \end{bmatrix} \begin{bmatrix} \mu_{G \cdot G} & \mu_{G \cdot K} \\ \mu_{K \cdot G} & v_{K \cdot K} \end{bmatrix} \begin{bmatrix} I_G & O \\ -v_{K \cdot K}^{-1} \mu_{K \cdot G} & I_K \end{bmatrix} = \begin{bmatrix} \lambda_{G \cdot G} & O \\ O & v_{K \cdot K} \end{bmatrix}$$

Since the matrix on the right-hand side of the equation is non-negative

definite, we have also

(35) $\quad \begin{bmatrix} \mu_{G \cdot G} & \mu_{G \cdot K} \\ \mu_{K \cdot G} & \mu_{K \cdot K} - \lambda_K \end{bmatrix}$ is non-negative definite.

Therefore λ_L is bounded by the inequality

(36) $\quad \lambda_L \geqq 0$

and by condition (35). We shall show that these bounds are the narrowest

possible if the distributions of z_I and w_I are normal and independent for

different values of t, and if condition (26) holds good. Let λ_L be any

diagonal matrix satisfying (35) and (36). We find $V_{K \cdot K}$ from the equation

$V_{K \cdot K} = \mu_{K \cdot K} - \lambda_K$. Since $V_{K \cdot K}$ is a principal minor of a non-negative

definite matrix by (35), it must itself be non-negative definite, and on account

of (26) it must be positive definite. Hence we can find a set z_K of normally

distributed variables which have $V_{K \cdot K}$ as covariance matrix. Next we determine

$\pi_{G \cdot K}$ from (31) and form a set of variables z_G by equation (14).

For the covariance matrix of z_I relations (29) and (30) must hold good,

so that $z_{I \cdot I}$ has the covariance matrix $V_{I \cdot I}$. Next we compute $\lambda_{G \cdot G} = \mu_{G \cdot G} - V_{G \cdot G}$

From (34) follows that $\lambda_{G \cdot G}$ is non-negative definite. We can now find a

set of normally distributed variables w_H which taken together with z_I has the

covariance matrix

$$\begin{bmatrix} V_{I \cdot I} & 0 & 0 \\ 0 & \lambda_{G \cdot G} & 0 \\ 0 & 0 & \lambda_L \end{bmatrix}$$

Finally we form the variables $x_H = z_H + w_H$ and $x = z$. Evidently x_I has $\mu_{I \cdot I}$ as

covariance matrix. This proves that our choice of λ_L corresponds to a possible

structure.

The region where condition (35) holds good is bounded by the sheet of

the hypersurface.

(37) $\quad \begin{vmatrix} \mu_{G \cdot G} & \mu_{G \cdot K} \\ \mu_{K \cdot G} & \mu_{K \cdot K} - \lambda_K \end{vmatrix} = 0$

which is nearest to origin in the positive orthant (i.e. the orthant where all coordinates are positive) in λ_L-space.

From (31) follows that $\pi_{g \cdot K}$ is a unique and continuous function of λ_L. Hence $\pi_{g \cdot K}$ will be confined to a region whose boundaries correspond to the boundaries of the admissible region for λ_L. In order to find the boundaries in $\pi_{g \cdot K}$-space, we use the expressions

$$(38) \quad \mu_{g k} + \pi_{g \cdot K} \mu_{K \cdot k} = \pi_{g k} \lambda_k$$

obtained from (30). The equation $\lambda_v = 0$ gives the n_v-flat

$$(39) \quad \mu_{g \cdot v} + \pi_{g \cdot K} \mu_{\cdot K \cdot v} = 0$$

in which the point $\pi_{g \cdot K}$ must lie.

From (36) and (38) we obtain that the admissible region in $\pi_{g \cdot K}$-space is bounded by the hyperplanes

$$(40) \quad \mu_{g L} + \pi_{g \cdot K} \mu_{K' \cdot L} = 0$$

These hyperplanes are identical with the bounding hyperplanes of the polytope $\beta_{g L \cdot K}^{(g K)}$ except for the hyperplane

$$(41) \quad \mu_{g \cdot g} + \pi_{g \cdot K} \mu_{K \cdot g} = 0$$

which is a bounding hyperplane of the polytope $\beta_{g L \cdot K}^{(g K)}$ but which does not occur in (40). The hyperplane (41) is now replaced by a sheet of the hypersurface in $\pi_{g \cdot K}$-space which we obtain when inserting the value of λ_L found from (38) in equation (37). This equation will be
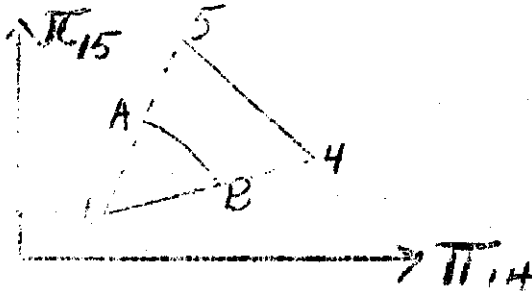
$$(42) \quad \begin{vmatrix} \mu_{G G} & \mu_{G \cdot L} & \mu_{G \cdot v} \\ \pi_L^{(g)} \mu_{L G} & \pi_L^{(g)} \mu_{L \cdot L} - \Delta_L & \pi_L^{(g)} \mu_{L \cdot v} \\ \mu_{v \cdot G} & \mu_{v \cdot L} & \mu_{v \cdot v} \end{vmatrix} = 0$$

where $\pi_L^{(g)}$ is a diagonal matrix containing the elements of the vector $\pi_{g \cdot L}$ and Δ_L is a diagonal matrix containing the elements of the vector $\mu_{g \cdot L} + \pi_{g \cdot K} \mu_{K \cdot L}$

The sheet of the hypersurface (42) which we want is the one which is nearest to the point $\beta_{g \cdot K}^{(g K)}$.

Example: Let $G = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$; $L = \begin{bmatrix} 4 & 5 \end{bmatrix}$, V empty.

The triangle (145) is the elementary regression triangle $L_{145\cdot45}^{(145)}$. The arc AB is a branch of the curve (42).

We shall now sum up our results in two theorems.

Theorem 3: <u>If there exists a g' such that the matrix $\rho_{g'L*}^{(g'K)}$ has compatible signs, then the structural point $\pi_{g K}$ is confined to that part of the polytope $\beta_{g L K}^{(g K)}$ which contains the vertex $\beta_{g K}^{(g K)}$ and which is cut off by the sheet of the hypersurface (42) which is nearest to this vertex.</u>

Theorem 4: <u>If z_t and w_t are normally distributed and independent for different values of t then the bounds given by Theorem 3 are the narrowest possible under the given assumptions.</u>

These theorems give bounds for the coefficient vectors of the equations of the reduced form. The results of course apply to any equation in the original system which contains only one post-determined variable. It is also possible to find bounds for a row vector $\alpha_{g\cdot L}$ of the matrix $\alpha_{g\cdot L}$ which is characterized by $n_g - 1$ zeros which may partly or wholly belong to $\alpha_{g K}$. We shall not here go into the determination of these bounds. We shall remark, however, that a necessary condition for the "surroundability" of a point $\alpha_{g L}$ by a bounded region is that at least $n_g - 1$ coefficients are known to be zero.

If there are more than $n_g - 1$ coefficients in each equation which are known to be zero, or if there are other additional linear relations, we may find the admissible region by first finding the admissible region when only the knowledge of $n_g - 1$ zeros are assumed and after that find the intersection between this region and the hypersurfaces corresponding to the other linear restrictions. Additional linear restrictions in the same equation will of course give hyperplanes, but additional linear restrictions in another

7. **The case when each equation contains only one post-determined variable and the covariance matrices of u_G and v_I are diagonal.**

In this case the original system is identical with the reduced form. Now $\alpha_{GG} = I$ and equation (12) takes the form

$$(43) \quad w_G = v_G + u_G$$

The matrix λ_{II} now is the sum of two diagonal matrices and must therefore also be diagonal. This means that we have now come back to the assumptions of non-correlation of Frisch (Ref. [2], Section 7) for the whole set u_I.

In this section we shall not presuppose any of the $\lambda's$ to be zero. This model also has much in common with the model underlying Thurstone's factor analysis. Our equation system may be written

$$(44) \quad z'_I = \begin{bmatrix} -\alpha_G \cdot K \\ I_K \end{bmatrix} z'_K$$

so that we may regard α_{GK} as part of a factor matrix. One difference between Thurstone's models and this model is that in our case certain prescribed coefficients of the factor matrix are supposed to be zero (more precisely, we suppose that the factor matrix is identical with its reduced form with respect to the set of variables K), while Thurstone supposes the factor matrix to contain a certain minimum of zeros without prescribing what individual coefficients are zero. My discussion of the identifiability of the uniquenesses (corresponding to λ_I in this paper) for given value of the number of common factors (corresponding to v_R in this paper) which I have given in a previously mimeographed Cowles Commission paper, applies to this case. Again there is a difference between our use of confluence analysis and factor analysis, in the respect that in factor analysis we are only interested in cases where the factor matrix is identifiable, preferably overidentified, while in confluence analysis we are also interested in cases where it is only possible to give bounds for the parameter points.

A consideration of this latter problem has been given in Ref. $\begin{bmatrix}6\end{bmatrix}$, Sections 9 - 11.

Again, in both cases we may test an hypothesis about the rank of the system. (i.e. the number of common factors in factor analysis and the number of pre-determined variables in confluence analysis). Again there is the difference that in confluence analysis we have a definite idea in advance not only about the number of important pre-determined variables but also what variables they are, while in factor analysis we have no definite idea about the number of common factors.

8. <u>Testing of the number of linear relations in a set of variables, and judging what variables should be included in a system of equations.</u>
In this section we shall assume for convenience that all variables x_i are normalized, i.e. that they are measured in units of their standard deviations.

In Monograph No. 10 of the Cowles Commission, Haavelmo has discussed a particular point in Frisch's confluence analysis. Frisch considers bunches of beams representing regression coefficients between normalized variables. According to Frisch, tight bunches show that there exists only one linear relation between the structural parts of the variables. This is true in the pure error model considered by Frisch and may be put in more exact form. If any of the bunches of a set of variables is confined to one quadrant then there cannot be more than one linear relation between the structural parts of this set. (This follows from Theorem 14 in Ref. $\begin{bmatrix}6\end{bmatrix}$). Haavelmo shows by an example that this is no longer true in a shock model. We may have two different linear relations and still have tight bunches. The testing of the number of linear equations has, however, never been a point of primary importance in confluence analysis.

It has been considered important mostly because the existence of two different linear equations between the structural parts in a pure error model destroys the possibility of determining bounds for any of them separately provided that we do not make any assumption about zero coefficients. In

Section 6 of this paper we have under certain conditions found bounds for the coefficient vectors of equations in the shock-error model. In order to find these bounds we have to find out in advance which variables are post-determined and which are pre-determined. But we need not know in advance if there are linear equations outside the investigational system. If there are outside linear equations which destroy the possibility of determining fairly narrow bounds for the structural vectors of the investigational system, this will show up in the analysis of the data, either in the way that the condition about compatibility of signs is not fulfilled or in the way that the admissible regions become too wide. Whether this will be the case or not will depend on the size of the variance of the sum $\alpha_j \sum_{j-1} x_{j-1} - u_j$. The larger this variance is compared with the variances of the disturbances of the investigational system the less influence will the j-th equation (which is supposed to be linear) have on the accuracy with which equations of the investigational system can be determined.

The whole question may, however, be put the other way round. Instead of asking how many linear equations there are in a given set of variables, we may ask how many variables, or rather what variables, we may include in the statistical determination of a given equation without destroying the reliability of the statistically determined coefficients. This is the important viewpoint in Frisch's theoretical publication on confluence analysis and still more in the practical applications of it. And this view-point applies very well also to the shock-error model with several post-determined variables. There will almost always be more variables which we suppose occur in a certain system than we can possibly include in the statistical analysis. If we include too few variables we have the risk of excluding variables whose influence is not negligible and which may be correlated with the pre-determined variables so that

the assumptions of our model do not hold good. If on the other hand we include too many variables the bounding regions become too large. If the bounding region of the parameter point becomes smaller when we include a new variable it is evident that it ought to be included. If it becomes a bit larger we may be in doubt whether to include the new variable or not. This procedure is of course no testing of an hypothesis, and we need not ask for the reason why the bounding region becomes larger when we include a new variable. This phenomenon will usually be a joint effect of the extent to which the new variable is correlated with the other pre-determined variables and the variance of the disturbance introduced by the new variable. (Compare Ref. [2], Section 14.)

To what extent it is due to one or the other of these causes need not bother us in the decision of whether to include the new variable or not.

REFERENCES:

1 : Cowles Commission Monograph No. 10

2 : R. Frisch; Statistical Confluence Analysis. Oslo, 1934

3 : R. C. Geary, Proceedings of the Royal Irish Academy, Vol. 47, Section A,

    No. 6 (1942)

4 : Ibid, Proceedings of the Royal Irish Academy, Vol. 49, Section A, No. 10 (1943)

5 : T. Koopmans; Linear Regression Analysis of Economic Time Series. 1937

6 : O. Reiersol; Econometrica, Vol. 9, p. 1 (1941)

7 : Ibid, Arkiv for Matematik, Astronomi och Tysik, Vol. 32A, No. 4 (1945)

8 : G. Tintner; Econometrica, Vol. 12, p. 97 (1944)

9 : Ibid, The Annals of Mathematical Statistics, Vol. 16, p. 304 (1945)

10 : A. Wald; Annals of Mathematical Statistics, Vol. 11, p. 284 (1940)