I. Introduction

1. Notation and terminology.

We shall introduce the following notations for sequences of numbers:

\[ J = \begin{bmatrix} 1 & 2 & \cdots & n_j \end{bmatrix} \]
\[ I = \begin{bmatrix} 1 & 2 & \cdots & n_i \end{bmatrix} \]
\[ G = \begin{bmatrix} 1 & 2 & \cdots & n_g \end{bmatrix} \]

We shall suppose that \( n_j \geq n_i \geq n_g \), 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 \( S' = [s_1 \ s_2 \ \cdots \ s_{n'}] \) and \( S'' = [s_1' \ s_2' \ \cdots \ s_{n''}'] \), the sequence \( [s_1' \ \cdots \ s_{n'}' \ s_1'' \ \cdots \ s_{n''}''] \) 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

\[ x_G = \begin{bmatrix} x_1 & x_2 & \cdots & x_{ng} \end{bmatrix} \]

A similar notation will be used for matrices, for instance

\[ \alpha_{G1} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n_1} \\ \cdots & \cdots & \cdots \\ \alpha_{ng_1} & \cdots & \alpha_{ngn_1} \end{bmatrix} \]

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
Let \( \mathbf{\eta}^{(y)} \) denote a vector function of \( y_i \) with \( n_y \) components, and let \( \mathbf{\gamma}^{(t)} \) denote the vector \( y_i \) at time \( t \). We shall suppose that \( \mathbf{\gamma}^{(t)} \) satisfies the following equations:

\[
(2) \quad \mathbf{\dot{\eta}}^{(t)}(y_t : I, y_{t-1} : I, \ldots, y_{t-n_y} : I) = \mathbf{u}_{\mathbf{\xi}}^{(t)}
\]

and

\[
(3) \quad \mathbf{\dot{\eta}}^{(t)}(y_t : \mathbf{\gamma}^{(t)}, y_{t-1} : \mathbf{\gamma}^{(t)}, \ldots, y_{t-n_y} : \mathbf{\gamma}^{(t)}) = \mathbf{u}_{\mathbf{\xi}}^{(t)}
\]

Equations (2) and (3) will be supposed to be such that they can be solved uniquely in terms of \( \mathbf{\gamma}^{(t)} \) and \( \mathbf{\gamma}^{(t-n_y)} \) respectively. \( \mathbf{u}_{\mathbf{\xi}}^{(t)}, \mathbf{u}_{\mathbf{\xi}}^{(t-n_y)} \) and \( \mathbf{\gamma}^{(t)} \) 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 \( y_i \) occur only at time \( t \) in equation (2) and do not occur at all in equation (3). In this case we assume that

\[
(4) \quad \mathbf{u}_{\mathbf{\xi}}^{(t)} \text{ is independent of } \mathbf{u}_{\mathbf{\xi}}^{(t-n_y)} \text{ when } u = 0, 1, 2,
\]

Case B: At least one variable \( y_i \) 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

\[
(5) \quad \mathbf{u}_{\mathbf{\xi}}^{(t)} \text{ is independent of } \mathbf{u}_{\mathbf{\xi}}^{(t-n_y)} \text{ when } u = 1, 2, \ldots
\]

For both cases we have

\[
(6) \quad \mathbf{u}_{\mathbf{\xi}}^{(t)} \text{ is independent of all variables } y_j \text{ in equation (2) except } y_{t-n_y}^{(t)}
\]

We shall suppose that, at least for the time being, we are only interested in getting information about Equation (3). This vector equation will be called the investigational system of equations and the set of variables occurring in it will be called the investigational set of variables. System (3) will be called the outside system and the variables will be called outside variables. The variables will be called post-determined variables.
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 \( y_{t-u,1} \) as a new variable \( y_{t,1} \). We shall use the notation \( I \) for the sequence of integers from 1 to the total number of variables after this renumbering, 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:

\[
(7) \quad \alpha_{G,I} y_I = u_I
\]

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 assumption and assumption (4). We shall suppose that

\[ E(y_I) = E(x_I) = 0 \]

II. Reduction to a pure error model.

Let the reduced form of 7 be

\[
(8) \quad y_G^i = \beta_{G,K} y_K = \beta_{G,G} u_G
\]

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

\[
(9) \quad s_G^i = y_G^i - \beta_{G,G} u_G
\]

\[
(10) \quad s_K = y_K
\]

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

\[
(12) \quad w_G^i = v_G^i + \beta_{G,G} u_G
\]
and

\[(13) \quad w'_K = u'_K\]

From (8) and (9) we obtain

\[(14) \quad Z'_Q + \sum \xi_K Z'_K = 0\]

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

\[(15) \quad \alpha'_I + \beta'_I - 0\]

Suppose that we have another set of variables \(x_Q\) which are correlated with the structural parts \(x_I\) 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

\[(16) \quad \alpha'_I E'(x'_I \cdot x_Q) = 0\]

The vector \(x'_I\) will be identifiable if the rank of \(E(x'_I \cdot x_Q)\) is \(n_I - 1\). 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 \( E(w_1 x_0) \) 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 \( Q_1 \). In the case where all covariances \( E(w_1 x_0) \) are equal to zero, my theorems about bounds in Ref. [6] and Ref. [7] reduce to theorems about identifiability of the vector \( Q_1 \) 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 \( u_0 \) independent of \( u_{j-q} \), \( u_0 \) may be supposed to be independent of all such variables. If we also assume the error \( v_1 \) to be independent of the error \( v_{j-1} \), then \( w_1 \) will be independent of \( x_{j-1} \). If the set \( x_{j-I} \) contains lagged values of the variables \( x_1 \), this leads us to introduce lagged values of the investigational variables as instrumental variables. We may also use lagged values of the variables \( x_1 \) 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.

## 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:

\[
\begin{align*}
\mu_{1-I} &= E(x'_{1-I} x_{1-I}) \\
\nu_{1-I} &= E(\frac{z^2_{1-I}}{\bar{z}_{1-I}}) E(z'_{1-I} z_{1-I})
\end{align*}
\]
\[ \lambda_{I,I} = E(\nu_{I}^{I}, \nu_{I}^{I}) \]

We shall make the following assumptions:

\[ E(\nu_{I}^{I}, \nu_{I}^{I}) = 0 \]
\[ E(\nu_{I}^{I}, \nu_{I}^{C}) = 0 \]
\[ E(\nu_{C}^{I}, \nu_{C}^{I}) = 0 \]
\[ E(\nu_{C}^{I}, \nu_{C}^{I}) \text{ is diagonal} \]
\[ \nu_{V} = 0 \text{ where } V \subseteq K \]

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 (17) and (18) follows that

\[ \mu_{I,I} = \nu_{I,I} + \lambda_{I,I} \]

From assumptions (18) - (20) follows that \( \lambda_{I,I} \) is of the form

\[ \lambda_{I,I} = \begin{bmatrix} \lambda_{C,C} & 0 & 0 \\ 0 & \lambda_{L} & 0 \\ 0 & 0 & 0 \end{bmatrix} \]

where we have written \( \lambda_{L} \) instead of \( \lambda_{L,L} \) 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 **polytope** is the extension to \( n \) dimensions of a polygon in 2 dimensions and a polyhedron in 3 dimensions.

An **orthant** is the extension to \( n \) dimensions of a quadrant in two dimensions and an octant in 3 dimensions.

A **p-flat** is a linear manifold of \( p \) dimensions.

A matrix \( \mu_{G,I} \) will be said to have **compatible signs** if there exist diagonal matrices \( E_{G} \) and \( E_{I} \) such that \( E_{G} \mu_{G,I} E_{I} \) 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

\[ (24) \quad \varepsilon_I + \pi_{KK} \varepsilon_K' = 0 \]

Let \( \hat{\mu}_{I,I} \) be the adjoint of \( \mu_{I,I} \), and let \( \beta_{I,I} \) be the matrix obtained from \( \hat{\mu}_{I,I} \) by dividing each row by the first element of that row. Let the vectors \( \beta_{I,K} \) and the vector \( \pi_K \) be interpreted as points in an \((n_1-1)\)-dimensional space. The convex polytopes with the \( n_h \) points \( \beta_{h,K} \) as vertices will be denoted by \( \beta_{h,K} \).

We have the following theorems:

Theorem I: If the matrix \( \beta_{h,K} \) has compatible signs, the structural point \( \pi_K \) is confined to the polytope \( \beta_{h,K} \).

Theorem II: If \( \varepsilon_I \) and \( \varepsilon_K' \) are normally distributed and if \( \varepsilon_I + \pi_{KK} \varepsilon_K' \) is independent of \( \varepsilon_I' + \pi_{IK} \varepsilon_K' \) when \( t \neq t' \), then each point in the polytope represents a possible structure under the given assumptions and the given probability distribution of \( \varepsilon_I \).

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 \( \mu_{I,I} \) and the conditions of the theorem.

All these statements follow immediately from Theorem 17.5 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 \( \pi_{KK} \). We shall consider one particular equation

\[ (25) \quad \varepsilon_I + \pi_{I,K} \varepsilon_K' = 0 \]
If we consider the set of variables \( y_{k} \), the matrix \( \lambda_{y} \) is diagonal so that Theorem I may be applied. Let us write \( y_{k} \) instead of \( y_{k} \cdot y_{k} \) and let
\[
\beta_{y_{k}}
\]
be the matrix which is obtained from \( \lambda_{y} \) after dividing each row by the first element of this row. If the matrix \( \beta_{y_{k}} \) has compatible signs, the structural point \( \lambda_{y} \) is confined to the polytope \( B_{y_{k}} \). 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 \( \lambda_{y} \) such that \( y_{k} \) is singular.

This assumption will certainly hold good if

(27) there exists at least one \( y \) such \( \beta_{y_{k}} \) has compatible signs.

This follows from a generalisation of Theorems 12 and 13 in Ref. [3].

Let us now again consider the whole reduced form (14). Postmultiplying by \( Z_{I} \) and taking the mathematical expectation, we obtain

\[
(28) \quad V_{G} \cdot I + \pi_{G} \cdot K \cdot V_{K} \cdot I = 0
\]

which may be split into two equations

\[
(29) \quad V_{G} \cdot G + \pi_{G} \cdot K \cdot \mu_{K} \cdot G
\]

and

\[
(30) \quad \mu_{G} \cdot K + \pi_{G} \cdot K \cdot V_{K} \cdot K = 0
\]

From (29) and (30) we obtain

\[
(31) \quad \pi_{G} \cdot K = -\mu_{G} \cdot K \cdot V_{K} \cdot K
\]

\[
(32) \quad V_{G} \cdot G = \mu_{G} \cdot K \cdot V_{K} \cdot K \cdot \mu_{K} \cdot G
\]

and

\[
(33) \quad \lambda_{G} \cdot G = \mu_{G} \cdot G - \mu_{G} \cdot K \cdot V_{K} \cdot K \cdot \mu_{K} \cdot G
\]

Using (33) we obtain the following identity

\[
(34) \quad I_{G} \cdot V_{K} \cdot K \cdot V_{K} \cdot K \cdot I_{G} = \begin{bmatrix} \lambda_{G} \cdot G & 0 & 0 \\ 0 & V_{K} \cdot K \end{bmatrix}
\]
Since the matrix on the right-hand side of the equation is non-negative definite, we have also

\[
\begin{bmatrix}
\mu_{G, G} & \mu_{G, K} \\
\mu_{K, G} & \mu_{K, K} - \lambda_K
\end{bmatrix}
\]

is non-negative definite.

Therefore \( \lambda_L \) is bounded by the inequality

\[
(56) \quad \lambda_L \geq 0
\]

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 (28) holds good. Let \( \lambda_L \) be any diagonal matrix satisfying (35) and (38). We find \( V_{K, K} \) from the equation

\[
V_{K, K} = \mu_{K, K} - \lambda_K.
\]

Since \( V_{K, K} \) 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 \( z_G \) of normally distributed variables which have \( V_{K, K} \) as covariance matrix. Next we determine \( \lambda_{G, K} \) from (31) and form a set of variables \( \bar{z}_G \) by equation (14).

For the covariance matrix of \( z_I \), relations (29) and (30) must hold good, so that \( \mu_I \) has the covariance matrix \( V_{I, I} \). Next we compute

\[
\lambda_{G, G} = \mu_{G, G} - V_{G, G}
\]

From (34) follows that \( \lambda_{G, G} \) is non-negative definite. We can now find a set of normally distributed variables \( z_I \) which taken together with \( \bar{z}_I \) has the covariance matrix

\[
\begin{bmatrix}
V_{I, I} & 0 & 0 \\
0 & \lambda_{G, G} & 0 \\
0 & 0 & \lambda_L
\end{bmatrix}
\]

Finally we form the variables \( \bar{z}_I \) and \( z_I \) as \( \lambda \). Evidently \( V \) has \( \mu_{I, I} \) as covariance matrix. This proves that our choice of \( \lambda \) corresponds to a possible structure.

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

\[
(37) \quad \begin{bmatrix}
\mu_{G, G} & \mu_{G, K} \\
\mu_{K, G} & \mu_{K, K} - \lambda_K
\end{bmatrix} = 0
\]
which is nearest to origin in the positive orthant (i.e. the orthant where all coordinates are positive) in $\lambda_L$-space.

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

$$\mu_g + \Pi_{g,K} \mu_K = \Pi_{g,K} \lambda_L$$

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

$$\mu_g + \Pi_{g,K} \mu_K = 0$$

in which the point $\Pi_{g,K}$ must lie.

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

$$\mu_g + \Pi_{g,K} \mu_K = 0$$

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

$$\mu_g + \Pi_{g,K} \mu_K = 0$$

which is a bounding hyperplane of the polytope $\beta^{(g,K)}$ but which does not occur in (40). The hyperplane (41) is now replaced by a sheet of the hypersurface in $\Pi_{g,K}$-space which we obtain when inserting the value of $\lambda_L$ found from (38) in equation (37). This equation will be

$$\begin{vmatrix}
\mu_G & \mu_{C,L} & \mu_{C,L} & \mu_{C,L} \\
\Pi_{g,K} \mu_{C,L} & \Pi_{g,K} \mu_{C,L} - A_L & \Pi_{g,K} \mu_{C,L} \\
\mu_{C,L} & \mu_{C,L} & \mu_{C,L} & \mu_{C,L} \\
\mu_{C,L} & \mu_{C,L} & \mu_{C,L} & \mu_{C,L}
\end{vmatrix} = 0$$

where $\Pi_{g,K}$ is a diagonal matrix containing the elements of the vector $\Pi_{g,K}$ and $A_L$ is a diagonal matrix containing the elements of the vector $\mu_{g,K}$.

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

Example: Let $G = [1, 2, 3]$; $L = [4, 5]$; $V$ empty.
The triangle (145) is the elementary regression triangle $\Delta_{45}$. The arc AB is a branch of the curve (42).

We shall now sum up our results in two theorems.

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

**Theorem 4:** If $\varepsilon_1$ and $\omega_1$ 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.

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 pre-determined variable. It is also possible to find bounds for a row vector $\alpha_{g^*, l}$ of the matrix $\alpha_{g^*, l}$ which is characterized by $\alpha_{g^*, l}$ zeros which may partly or wholly belong to $\alpha_{g^*, l}$. We shall not here go into the determination of these bounds. We shall remark, however, that a necessary condition for the "surrondability" of a point $\alpha_{g^*, l}$ by a bounded region is that at least $\alpha_{g^*, l}$ coefficients are known to be zero.

If there are more than $\alpha_{g^*, l}$ 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 $\alpha_{g^*, l}$ 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 $\nu^c$ and $\nu^T$ are diagonal.

In this case the original system is identical with the reduced form. Now $\alpha^G G^G = I$ and equation (12) takes the form

$$ (45) \quad \nu^G = \nu^0 + \nu^G. $$

The matrix $\lambda^I$ 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 $\nu^G$.

In this section we shall not presuppose any of the $\lambda^I$ to be zero.

This model also has much in common with the model underlying Thurstone's factor analysis. Our equation system may be written

$$ (44) \quad Z^I = \begin{bmatrix} -\alpha^G I^K \\ I^K \end{bmatrix} Z^K $$

so that we may regard $\nu^G$ 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 $\lambda^I$ in this paper) for given value of the number of common factors (corresponding to $\nu^G$ 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 overdetermined, 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. [6], 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. Testing of the number of linear relations in a set of variables, and judging what variables should be included in a system of equations.

In this section we shall assume for convenience that all variables \( \mu \) 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. [6]). 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 $\sum_j \left( x_j - \bar{x}_j \right)^2$. 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 viewpoint 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 negli-
gible 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 16.)

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, 1954
3. R. C. Geary, Proceedings of the Royal Irish Academy, Vol. 47, Section A, No. 6 (1942)
5. T. Koopmans; Linear Regression Analysis of Economic Time Series. 1937
6. C. Reiersol; Econometrica, Vol. 9, p. 1 (1941)
8. G. Tintner; Econometrica, Vol. 12, p. 97 (1944)