Discussion Papers
Department of Economics
University of Copenhagen

No. 12-01

Haavelmo's Probability Approach and the Cointegrated VAR

Katarina Juselius

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Haavelmo's Probability Approach and the Cointegrated VAR*

Katarina Juselius
University of Copenhagen, Department of Economics

March 1, 2012

Abstract

Some key econometric concepts and problems addressed by Trygve Haavelmo and Ragnar Frisch are discussed within the general framework of a cointegrated VAR. The focus is on problems typical of time-series data such as multicollinearity, spurious correlation and regression results, time dependent residuals, normalization, reduced rank, model selection, missing variables, simultaneity, autonomy and identification. Specifically the paper discusses (1) the conditions under which the VAR model represents a full probability formulation of a sample of time-series observations, (2) the plausibility of the multivariate normality assumption underlying the VAR, (3) cointegration as a solution to the problem of spurious correlation and multicollinearity when data contain deterministic and stochastic trends, (4) the existence of a universe, (5) the association between Frisch's confluence analysis and cointegrated VAR analysis, (6) simultaneity and identification when data are nonstationary, (7) conditions under which identified cointegration relations can be considered structural or autonomous, and finally (8) a formulation of a design of experiment for passive observations based on theory consistent CVAR scenarios illustrated with a monetary model for inflation.

Keywords: Haavelmo, CVAR, autonomy, identification, passive observations

JEL classification: B16, B31, B41, C32, C82

*Valuable comments from Kevin Hoover and Roman Frydman are gratefully acknowledged.

1 Introduction

Haavelmo's Nobel prize winning Econometrica (1944) monograph "The Probability Approach to Econometrics" is widely considered to have laid the foundations for modern macro-econometrics. See for example the survey articles in Aldrich (1989), Hendry and Morgan (1989), Hendry, Spanos, and Ericsson (1989), Anderson (1992) and the edited volume "The Foundations of Econometric Analysis" by Hendry and Morgan (1995).

In today's econometric world, new concepts, tests and estimators are developed side by side with empirical applications. This was less so in Haavelmo's time when empirical analyses which now can be done within seconds would have required years of work, if at all possible. Given these obstacles, his vision about econometrics based on time series data by "passive observation" is truly remarkable.

At this background the idea of the paper is to discuss a number of econometric problems such as multicollinearity, spurious correlation and regression, time dependent residuals, normalization, reduced rank, model selection, missing variables, simultaneity, autonomy and identification. While these problems were fairly well understood at the time, they were nonetheless subject to considerable difficulties on a practical level. The idea of this paper is, therefore, to address the question whether we now have found a practical solution to these problems.

A distinguishing feature of the econometrics in Haavelmo's time compared to our's is the advancement of the theory of integrated processes. The Cointegrated VAR (CVAR) model (Johansen, 1996, Juselius, 2006) offers a practical methodology for analyzing such nonstationary data in a maximum likelihood framework. Likelihood inference based on a joint probability formulation of the observables is one of Haavelmo's most important contributions and Section 2 will discuss in what sense the CVAR represents Haavelmo's vision of a such an approach. Section 3 and 4 will then discuss solutions to the above econometric problems within a general CVAR framework while recognizing that other econometric approaches may offer equally good solutions.

Haavelmo's work was strongly influenced by Ragnar Frisch and it is difficult to ignore Frisch when discussing the above issues. In the words of Aldrich (1989, p.1): "The concepts relating to structure were devised by Frisch and they passed into classical econometrics through Trygve Haavelmo's Probability Approach in Econometrics (1944). At the centre of Frisch's network of concepts was 'autonomy', or, as it is now more often called, 'structural invariance'." Thus, many of the basic concepts first formulated by Frisch were later reformulated by Haavelmo within his probability approach. The fact that Frisch was not fully convinced that a joint probability formulation is a solution to the above econometric problems makes it even more interesting to re-address them here in our chosen probability framework.

One of Haavelmo's important methodological contributions was to distinguish between theoretical, true and observed variables:

> We may express the difference by saying that the "true" variables (or time functions) represent our ideal as to accurate measurements of reality "as it is in fact," while the variables defined in a theory are the true measurements that we should make if reality were actually in accordance with our theoretical model. [Haavelmo, 1944, p. 5]

and

> One could perhaps also characterize the difference between the "true" and the "observational" variables in the following way. The "true" variables are variables such that, if their behavior should contradict a theory, the theory would be rejected as false; while "observational" variables, when contradicting the theory, leave the possibility that we might be trying out the theory on facts for which the theory was not meant to hold, the confusion being caused by the use of the same names for quantities that are actually different. [Haavelmo, 1944 p. 7]

The quotation marks around "true" and reality "as it is in fact " seem to signal that he considered the link between the true variables suggested by the theory and the actual measurements taken from official statistics to be very weak. Definitions often change, new components frequently enter the aggregates and new regulations change the economic behavior and high quality, reasonably long aggregate series are notoriously difficult to find. But, as these data are the best set of measurements in the circumstances, Haavelmo argued that we better adjust our theories and econometric methodology to fit this reality "as it is" rather than the other way around.

To avoid "a good deal of planless and futile juggling with figures" he argued that a selected theory should always be accompanied by a design of experiment that "describes and explains how to measure the variables supposed to be the true variables". To further emphasize the distinction between testing a theory in an ideal (but often unrealistic) situation and a less ideal (but realistic) situation he discussed two kinds of experimental situations in which the theoretical model is to be taken to the data:

> If economists would describe the experiments they have in mind when they construct the theories they would see that the experiments they have in mind may be grouped into two different classes namely, (1) experiments that we should like to make to see if certain real economic phenomena -when artificially isolated from other influences -would verify certain hypothesis and (2) the stream of experiments that nature is steadily turning out from his own enormous laboratory, and which we merely watch as passive observers. [Haavelmo, 1944, p. 8].

In the first case, it is straightforward to formulate an adequate probability model and test the theory against the measurements of the 'true' variables. In the other case, it is much more demanding to take a theory to data without violating scientific principles. One of Haavelmo's important contributions is to have provided us with a coherent analytical framework for the analysis of data based on passive observations where the notion of a design of experiment, addressed in Section 5, plays a crucial role for improving the fit between theory and evidence. Finally, Section 6 provides an illustration of how Haavelmo's idea of a design of experiment for passive observations can be formalized in the form of a theory consistent CVAR scenario which translates the basic assumptions of a theory model into testable hypotheses on the 'pulling and pushing' forces of the model. The paper ends with a discussion of how a joint probability formulation of all observables combined with a theory consistent scenario provides information on how to adjust the theory, "so as to make the facts we consider the "true" variables relevant to the theory".

2 Haavelmo's probability approach

Consider a time series of p variables motivated by an interest in a theoretical macroeconomic relationship, $x_{1,t} = f(x_{2,t}, \cdots, x_{p,t})$, $t = 1, .., T$. At each point in time, t, there is just one realization, $x_t' = [x_{1,t}, x_{2,t}, \cdots, x_{p,t}]$ of the underlying stochastic process. The full sample of observations is given by:

$$X = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{p,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{p,2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1,T} & x_{2,T} & \cdots & x_{p,T} \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_T' \end{bmatrix}$$

In Haavelmo's time, an economic relation was in most cases specified as a linear regression model and estimated by ordinary least squares. However, both Haavelmo and Frisch were concerned about the fact that successive observations of typical macroeconomic variables were highly dependent rendering the ordinary regression model less suitable. Haavelmo's solution was to formulate the joint probability, $P(X; \theta)$, of the sample point, X, and then estimate the parameters θ based on maximum likelihood rather than OLS.

> ... it has been argued, e.g., that most economic time series do not conform well to any probability model, "because the successive observations are not independent." But it is not necessary that the observations should be independent and that they should all follow the same one-dimensional probability law. It is sufficient to assume that the whole set of, say n, observations [where n means $T \times p$ above] may be considered as one observation of n variables (or a "sample point") following an n-dimensional joint probability law, the "existence" of which may be purely hypothetical. Then, one can test hypotheses regarding this joint probability law, and draw inference as to its possible form, by means of one sample point (in n dimensions). [Haavelmo, 1944, Preface, *iii*]

The next section will discuss the conditions under which the VAR model can be considered a suitable candidate for this kind of time series data.

2.1 Deriving a baseline VAR

Consider the joint probability of X given the initial value X_0 with the parameter value θ:

$$P(X|X_0; \theta) = P(x_1, x_2, ..., x_T|X_0; \theta).$$

If the multivariate normal distribution is a reasonable approximation to the probability law, P, then $\theta = f(\mu, \Sigma)$ where μ is the mean and Σ the covariance matrix, Σ. It is convenient to express the joint probability of $(X|X_0)$ as the probability of the stacked process $Z' = [x_1', x_2', x_3', ..., x_T'] \sim N_{Tp}(\mu, \Sigma)$. Since μ is $Tp \times 1$ and Σ is $Tp \times Tp$, there are many more parameters than observations and it is not possible to get unique estimates of μ, Σ without making simplifying assumptions, for example that the covariances, $E(x_t x_t')$, are approximately constant over the sample period.

A more useful formulation can be obtained by decomposing the joint probability into a conditional and a marginal probability and then sequentially repeating the decomposition for the marginal probabilities (Hendry and Richard, 1983):

$$\begin{aligned} & P(x_1, x_2, x_3, ..., x_T | X_0; \theta) \\ = & P(x_T | x_{T-1}, ..., x_1, X_0; \theta) P(x_{T-1}, x_{T-2}, ..., x_1 | X_0; \theta) \\ & \vdots \\ = & \prod_{t=1}^{T} P(x_t | X_{t-1}^0; \theta) \end{aligned} \tag{1}$$

where

$$X_{t-1}^0 = [x_{t-1}, x_{t-2}, ..., x_1, X_0].$$

By assuming that the probability distribution follows the multinormal law, the conditional model becomes:

$$(x_t | X_{t-1}^0) \sim N(\mu_t, \Sigma_{11.2})$$

where

$$\mu_t = m_1 + \Sigma_{12} \Sigma_{22}^{-1} (X_{t-1}^0 - m_2) \tag{2}$$

and

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \tag{3}$$

The difference between the observed value of the process and its conditional mean is denoted ε_t :

$$x_t - \mu_t = \varepsilon_t$$

Inserting the expression for the conditional mean gives:

$$x_t = m_1 - \Sigma_{12} \Sigma_{22}^{-1} m_2 + \Sigma_{12} \Sigma_{22}^{-1} X_{t-1}^0 + \varepsilon_t.$$

Using the notation: $\mu_0 = m_1 - \Sigma_{12}\Sigma_{22}^{-1}m_2$, $[\Pi_{1,}\Pi_2, ..., \Pi_{T-1}] = \Sigma_{12}\Sigma_{22}^{-1}$ and assuming that $\Pi_{k+1}, \Pi_{k+2}, ..., \Pi_{T-1} \simeq 0$, we arrive at the k^{th} order vector autoregressive (VAR) model:

$$x_t = \mu_0 + \Pi_1 x_{t-1} + \cdots + \Pi_k x_{t-k} + \varepsilon_t,\ t = 1, ..., T \tag{4}$$

where ε_t is $NI_p(0, \Omega)$ and $x_0, ...x_{-k+1}$ are assumed fixed. Thus, even when the observations x_t are strongly time dependent, the conditional process $(x_t | X_{t-1}^0)$ is independent and OLS estimates of $\{\Pi_1, ..., \Pi_k, \mu_0, \Omega\}$ are Maximum Likelihood estimates.

Hence, the VAR model is essentially a reformulation of (time invariant) covariances of the data. In this sense it can be considered a first general approximation to the actual data generating process (Hendry and Mizon, 1993).

2.2 Are the assumptions underlying the VAR plausible?

The VAR model in (4) is based on the assumption that the residuals are multivariate normal which implies:

1. Parameter constancy: $\mu_0, \Pi_1, \cdots, \Pi_k, \Omega$ are constant over time, i.e. the covariances are constant over time.

2. Residual independence.

3. Valid truncation $(\Pi_{k+1}, \Pi_{k+2}, ..., \Pi_{T-1} \simeq 0)$.

Many economists tend to consider multivariate normality a convenient assumption that they would not expect to hold in economic data. This was also a view Frisch shared:

> Frisch's convictions about the structure of economic reality paired with 'passive observations' left him in no doubt that normality or other reasonable distributions were unlikely to be fulfilled, as required by standard method of statistical analysis. [Bjerkholt, 2011, p. 9]

and

> Frisch was here and in other projects where he hunted for alternative approaches, very skeptical about falling back on the least squares method and perhaps even more about making unwarranted assumptions about normality [Bjerkholt, 2011, p. 12].

Even though there is no *a priori* reason as such to expect the VAR residuals to be normally distributed, Haavelmo provided some arguments that can be used as a justification for the normality assumption:

> ... if we consider a set of related economic variables, it is, in general, not possible to express any one of the variables as an exact function of the other variables only. There will be an "unexplained rest," and, for statistical purposes, certain stochastic properties must be ascribed to this rest, a priori. Personally I think that economic theorists have, in general, paid too little attention to such stochastical formulation of economic theories. For the necessity of introducing "error terms" in economic relations is not merely a result of statistical errors of measurement. It is as much a result of the very nature of economic behavior, its dependence upon an enormous number of factors, as compared with those which we can account for, explicitly, in our theories. We need a stochastical formulation to make simplified relations elastic enough for applications. [Haavelmo, 1944, p. 1]

Thus, Haavelmo considers the residuals to be a catch-all for everything else that is not included in the empirical model. When this 'everything else' comprises an 'enormous number of factors' the central limit theorem would suggest that normality could be approximately valid, provided these factors are independent.

But, since we do not know if they are, normality is a possibility that needs to be checked against the data. Indeed, when the VAR model in is routinely estimated, the tests usually reject not just normality but also homoscedasticity and even independence. In most cases this is because the effect of extraordinary events related to changes in political and economic institutions, to reforms and interventions may not have been fully anticipated. If they were, we should not see these large effects in the variables in question.

Such events can be highly influential (Nielsen, 2008) and, unless adequately controlled for, tend to bias the estimates of the model parameters. In many cases we can use various dummies to control *ex post* for extraordinary events. For example, step dummies can be used to model equilibrium mean shifts (in $\beta' x_t$) and impulse dummies to account for extraordinary shocks in the equations (in Δx_t).[1] But since these dummies operate *ex post* they are not helpful for unconditional forecasting.

Failure to properly control for extraordinary events is likely to cause residuals to be autocorrelated. For example, a non-modeled equilibrium mean shift or a change in the growth rates will automatically produce residual autocorrelations and may (incorrectly) suggest longer lags in the VAR. When the breaks are properly modelled, excessive autocorrelation usually disappears and a correct lag length can be determined.[2]

Haavelmo was well aware of the possibility of such extraordinary events as appears from the following citation:

> Purely empirical investigations have taught us that certain things in the real world happen only very rarely, they are "miracles," while others are "usual events." The probability calculus has developed out of a desire to have a formal logical apparatus for dealing with such phenomena of real life. The question is not whether probabilities exist or not, but whether-if we proceed as if they existed-we are able to make statements about real phenomena that are "correct for practical purposes." [Haavelmo, 1944, p. 43]

The "usual events" can often be adequately described by a normal distribution, whereas the "miracles" tend to fall outside the normal range: they are the outlying events discussed above. The Haavelmo citation suggests that he either thought the "miracles" could be referred to the normal distribution, or that he had a different distribution in mind such as the Cauchy or t-distribution. For macroeconomic data the latter may not be the most obvious choice as the "miracles" are seldom random but rather a consequence of omitted institutional events. Of course, also the "usual events" can often be related to policy changes, interventions, etc., but as they are numerous in

[1] Note that in a dynamic regression model, like the VAR, a dummy variable takes controls for the unanticipated shock, but leaves the observation intact. This is contrary to a static regression model where a dummy variable removes the observation altogether.

[2] Most empirical VAR models do not control for breaks and, therefore, are prone to use too many lags.

numbers and often of minor importance they can for practical purposes be treated as random events.

Finally, the important question of parameter constancy needs to be addressed. In Chapter II, about the "degree of permanence of economic laws" Haavelmo raised the question "whether or not we might hope to find elements of invariance in economic life,upon which to establish permanent 'laws'":

> When we use the terms "constant relationships," or "unstable, changing relationships," we obviously refer to the behavior of some real economic phenomena, as compared with some behavior that we expect from theoretical considerations. The notion of constancy or permanence of a relationship is, therefore, not one of pure theory. It is a property of real phenomena as we look upon them from the point of view of a particular theory. [Haavelmo, 1944, p. 13]

While Haavelmo clearly emphasized that parameter constancy is a property of real phenomena rather than of a theoretical model, the actual testing for structural change would probably have been too time consuming to carry out in his time. In today's econometric world of fast computers, tests for parameter constancy and structural change abound.[3] Provided the sample is sufficiently large, numerous tests can be used to check the assumption of parameter constancy. Such acceptance depends, often crucially, on whether all major known structural changes and regime shifts have been adequately accounted for. If the covariance matrices Σ_{12} and Σ_{22} have changed (as a result of reforms or interventions during the sample period), then the VAR parameters will also exhibit time variation unless these events are properly controlled for. In this sense, the VAR generally corresponds to an *ex post* in-sample probability model with constant parameters (after correcting for changes in structure).

Ultimately the sample might have to be redefined to achieve an acceptable level of constancy if the change in regime was sufficiently profound. In this sense, the issue of structural change is crucially related to the existence of a universe (population) of true economic models that stay constant over time. Whether one could assume such a universe or not was of great concern to

[3]For example, Hansen and Johansen (1999) provide a battery of such tests tailormade for the cointegrated VAR model.

Frisch and his contemporaries, in particular Keynes.[4] The following conversation between Frisch and Koopmans recorded and interpreted in Bjerkholt (2011, p.12) gives some flavor of the debate:

> Koopmans means that the assumptions about distribution laws etc. must lead back to the assumption about a universe.
>
> Frisch means that we should build on the observations as they actually are and do our assumptions on the basis of "the sample as it is". We consider the variable as consisting of two parts, a systematic part and an error element. The assumptions about the error element are based on the actual example. "But I do not object to the idea of the universe, in many cases it may help. But elsewhere not, particularly in economic data of historical character, data which cannot be repeated."
>
> Koopmans: Does this framework have any meaning when the sample is small? For what shall we apply our empirically determined coefficients if we do not want to consider the concept of a universe?
>
> Frisch: When we from a certain set of observations have determined certain coefficients, we will naturally at the next opportunity try to determine the same coefficients again and compare the results. The true character of the universe cannot be observed.
>
> Koopmans: We can determine whether our observed sample is normally distributed without employing the concept of a universe. All characteristics are applied to the sample.

The question of a universe is still highly relevant today (in particular as conventional models based on a constant-parameter universe failed to foresee one of the deepest economic and financial crisis of our time) and will be discussed in more detail in Section 6.3 at the end of this paper.

3 Correlation, Cointegration, and Confluence Analysis

Much of the debate on econometric method was related to the regression model and its many shortcomings when applied to time-series data. That

[4] For a detailed discussion of Keynes on Tinbergen's econometric method see Garrone et al. (2004).

most economic variables are trending and subject to strong multicollinearity was something Frisch, in particular, was concerned about. The correlation coefficient would in this case be dominated by the trend (the long-time component) even though the aim was to study the relationship between the cyclical parts (the short-time components) of the variables. In a discussion of the Yale lectures, Bjerkholt (2011) argues that Frisch knew from the application of principal component theory how misleading a correlation coefficient of trending variables would be and, therefore, used it to show 'the necessity of decomposing our time series before analyzing their interrelationships".

3.1 Correlation contra cointegration analysis when data are trending

That correlation/regression coefficients for trending data are prone to produce nonsense correlations/regression coefficients was already demonstrated by Yule (1926). Phillips (1987) showed that unit roots in the data would produce such spurious results. The theory of unit root econometrics was not yet developed in Haavelmo's time and trends were mostly assumed to be deterministic but, as we know today, trends in economic variables are most stochastic. In such cases, the average value of a unit root variable is an inconsistent estimate of it mean and the empirical correlation coefficient is a meaningless estimate of the true association between two variables (Phillips, 1987). Thus, Frisch' concern about the misuse of correlation coefficients between economic variables was, therefore, relevant even when the variables were trend-adjusted.

The CVAR circumvents this problem by formulating the general VAR in the error-correction form:

$$\Delta x_t = \mu_0 + \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} \cdots + \Gamma_{k-1} \Delta x_{t-k+1} + \Phi D_t + \varepsilon_t,\ t = 1, ..., T \tag{5}$$

$$\varepsilon_t \sim NID(0, \Sigma). \tag{6}$$

where ΦD_t contains all deterministic components (trend, constant, dummies). The hypothesis that $x_t \sim I(1)$ is formulated as a reduced rank condition:

$$\Pi = \alpha \beta' \tag{7}$$

where α and β are $p \times r$ matrices ($r < p$) and the r relations, $\beta' x_t$, define stationary combination of nonstationary variables. Assuming just two lags for simplicity, the cointegrated VAR model becomes:

$$\begin{aligned} \Delta x_t &= \mu_0 + \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \Phi D_t + \varepsilon_t,\ t = 1, ..., T \\ \varepsilon_t &\sim NID(0, \Sigma). \end{aligned} \tag{8}$$

By transforming the trending variables, x_t, into stationary differences, Δx_t, and stationary cointegration relations, $\beta' x_t$, two of the problems discussed by Frisch are basically solved:

1. Multicollinearity between the $\Delta x's$ and $\beta' x_t$ has been reduced to a minimum by the removal of trends (stochastic as well as deterministic) achieved by differencing and cointegration.

2. Regression and correlation coefficients between the stationary components are well defined and, for given β, standard inference on $(\alpha, \Gamma_1, \Sigma)$ applies.

The model is nonlinear in α and β, but can be estimated by reduced rank regressions as shown in Johansen (1988) where the β relations are found as the eigenvectors to a solution of an eigenvalue problem and α is estimated by linear regression for a given β. The relations $\beta' x_t$ define r linear relationships between p variables. These seem closely related to Frisch' problem of 'identifying' an economic relation which belonged to a system of relations, where the the number of economic relations is smaller than the number of variables.

Two variables are cointegrated only if they share a common stochastic trend defined as the cumulation of all permanent shocks that have pushed the variables out of equilibrium. This is a demanding measure of association which avoids the pitfall of a spurious correlation coefficient. But, contrary to a correlation coefficient, a bivariate cointegration relation is empirically well defined only if the two variables share one common stochastic trend. As most economic variables share more than one stochastic trend, we usually need more than two variables to establish cointegration.

A cointegration relation is, therefore, more closely related to a multiple regression relation. But there is an important difference: The estimates of a cointegration relation are invariant to extensions of the variable set whereas the linear regression estimates are only invariant when regressors are orthogonal. Thus, if a cointegration relation has been found between a set of variables, the same relation will be found in a larger set of variables with identical coefficients up to sample variations.[5] Thus, the problem of multicollinearity which Frisch tried to solve with confluence analysis is not present in cointegration analysis.

3.2 Confluence analysis

Already in the late 1920's, Frisch and his contemporaries had become concerned about the problem of several relationships holding simultaneously in the data. The core question, which later became the question of identification, was how to unravel the relationships of interest from the ones which were a characteristic of the data set, but of no interest to the economist. (Hendry and Morgan, 1989).

Frisch was concerned about the fact that regression analysis could be very misleading when the relation of interest belonged to a system of relations, i.e. when there was "multicollinearity" between the economic variables. As a means to uncover multicollinearity problems in the data he suggested the method of 'confluence' analysis and 'bunch maps'. To see how confluence analysis relates to cointegration analysis it is useful to briefly discuss the former as "a diagnostic tool aimed at disentangling simultaneous relations, identifying different relations present, and related problems, including choices of regression methods" (Bjerkholt, 2011, p.6).

The following example is broadly based on Hendry and Morgan (1989) and describes a system of exact relations where the observed variables, x_t, are subject to measurement errors:

$$x_t = \xi_t + v_t$$

and ξ_t is a latent unobservable variable, v_t is a measurement error, $E(\xi_t' v_t) = 0$ and $E(v_t' v_t) = \Sigma$. There are $k \leq n$ constant linear relationships between ξ_t, $A\xi_t = 0$, where A is $k \times n$. Then $Ax_t = u_t$, where $u_t = Av_t$ and $E(x_t x_t') = M = \phi + \Sigma$.

Based on this model set-up, Hendry and Morgan (1989) identifies four issues which Frisch was concerned about:

1. Normalizing on one $x_{j,t}$ in each equation and regressing that variable on the remaining $x's$ will not consistently estimate the elements of A since

[5] Adding new variables to the model will, however, in most cases result in new cointegration relations.

$E(x_t u_t') \neq 0$. This is the measurement error or simultaneous equations inconsistency.

2. Since the rank of $\phi = n - k$, the moment matrix M would be singular without errors of measurement. Thus, regression estimates of A are determinate only because of observational errors (better data would make this worse!) and so Frisch regarded such estimates as being nonsense. This is the reduced rank problem.

3. Since $A\xi_t = 0$, $DA\xi_t = A^*\xi_t = 0$ for any nonsingular $k \times k$ matrix D. This is the problem of identification. Distinguishing A (the economic theory parameters of interest) from A^* requires imposing restrictions on A. This is the problem of identification.

4. Finally, if k and the coefficients in A are not known *a priori* they have to be selected from the data. This is the model selection problem.

Confluence analysis was designed to solve some of the obvious problems plaguing regression analysis, but like regression analysis it was non-structural (Aldrich, 1989), a view also held by Haavelmo. Hendry and Morgan (1989) give a number of reasons why they consider confluence analysis outmoded in today's econometrics world. They conclude their discussion by arguing that cointegration analysis essentially provides a solution for problems 1 and 2. The first, normalization, does not pose a problem in cointegration analysis as the (relative) size of a cointegration coefficient is canonical and does not change with normalization. Problem 2 is solved by finding the reduced rank of the long-run matrix Π. Problem 3, how to identify the underlying economic relations, will be discussed in the next section.

Problem 4, how to select a correct (or adequate) model, was discussed in terms of the missing-variable problem. Frisch tackled this by first decomposing the variations of a modeled variable into three parts: 'systematic variations', 'disturbances' and 'accidental variations'. The variations are 'systematic' when all the relevant explanatory variables are correctly included; 'disturbances' occur when one or a few highly significant variables are missing whereas 'accidental variations' are the result of omitting variables which are unimportant and negligible. (Bjerkholt, 2011)

Can the CVAR residuals be considered disturbances or accidental variations? Because basically everything depends on everything else in economics, the missing variable problem is hard to escape whatever the empirical approach taken. Most researchers decide for a selection of the key economic variables knowing that this will leave out relevant variables. As CVAR analyses which tend to become rather involved when the number of variables becomes large, most CVAR models are based on rather few variables. Thus, the missing variable problem is present also in this approach. But, by allowing for a sufficiently long lag structure the VAR model can to some extent compensate for missing variables so that the size of the residuals may not differ much from 'accidental variations'. Nonetheless, if we include previously omitted variables to the model some of the empirical conclusions may change. In particular, the short-run adjustment coefficients are not invariant to extensions in the variable set contrary to the long-run cointegration relations. For example, a variable found to be exogenous in a smaller model may no longer be so in the extended model. By gradually expanding the model the effect of the *ceteris paribus* assumption on the empirical conclusions can be exploited. For an illustration see Juselius (2006).

To conclude, the VAR model has provided a solution to the problem of (1) time dependent residuals by conditioning on sufficiently many lags, (2) spurious correlation and regression results, (3) multicollinearity, (4) normalization, (5) reduced rank, and (6) model selection to some extent. The question to be discuss next is whether the CVAR can also "unravel the relationships of interest from the ones which were a characteristic of the data set, but of no interest to the economist" i.e. whether the CVAR can also provide a solution to the problem of identification.

4 Simultaneity and identification

The interrelated issues, simultaneity, autonomy and identification, were of vital interest to Frisch and Haavelmo. Frisch proposed confluence analysis as a tool of uncover structure among correlated variables but in the end became quite sceptical of the possibility to identify and estimate structural models. Most of the applications of confluence analysis dealt with the problem of estimating a single equation with too many variables (Aldrich, 1989 p.2). Haavelmo, on the other hand, proposed a solution based on the formulation of a probability model for the data so that a generically identified system of equations could be estimated by maximum likelihood methods. In this sense he associated identification with the available data. This is also the spirit in which Johansen and Juselius (1994) discuss three aspects of identification :

1. *Generic identification* is needed to be able to uniquely estimate the parameters. Without generic identification it is not possible to solve the statistical model. Generic identification as such does not guarantee economic structure.

2. *Empirical identification* refers to the actual estimated parameter values and is needed to secure that a generically identified structure holds empirically. This condition needs to be satisfied because if an identifying coefficient is statistically insignificant and set to zero generic identification may break down.

3. *Economic identification* is related to the economic plausibility of the estimated coefficients of a generically and empirically identified relation/model. In particular, claiming that the generically and empirically identified relations are estimates of autonomous (structural) economic relations requires economic identification.

The first two criteria are uncontroversial and can easily be checked, whereas the third one is what Haavelmo and Frisch were primarily concerned about. The discussion below will center around this aspect and how it needs to be reformulated when data are nonstationary.

4.1 Identification when data are nonstationary

That unit root nonstationarity seems endemic in economic data has some important implications for how to address identification. Juselius (2006) discusses four identification problems when data contain unit roots: the identification of (1) the long-run cointegration relations, (2) the short-run adjustment structure, (3) the exogenous driving shocks, and (4) the dynamics of the impulse responses. For simplicity of exposition, only the first two will be discussed here. Of them, it is the identification of the long-run structure, $\beta' x_t$, that seems most closely related to Haavelmo's 1943 paper on the identification of of simultaneous equations. But, while the distinction between long-run and short-run is not present in the 'traditional' simultaneous equations system, this is not the case in the CVAR model.

To examine the relationship between long-run and short-run identification, it is useful to start with a CVAR model where current (simultaneous) effects are not explicitly modeled but left to the residual covariance matrix Ω:

$$\Delta x_t = \Gamma_1 \Delta x_{t-1} + \alpha\beta' x_{t-1} + \Phi D_t + \varepsilon_t, \quad \varepsilon_t \sim IN(0, \Omega) \tag{9}$$

Without changing the likelihood, one can pre-multiply (9) by a nonsingular $p \times p$ matrix A_0 representing (short-run) current effects that may or may not have a structural interpretation:

$$A_0 \Delta x_t = A_1 \Delta x_{t-1} + a_1 \beta' x_{t-1} + \tilde{\Phi} D_t + v_t, \quad v_t \sim IN(0, \Sigma). \tag{10}$$

It appears that β is the same in (9) and (10) and therefore can be estimated based on either form. Furthermore, the estimates of β are super consistent, i.e. the speed of convergence toward the true value β is proportional to T as $T \to \infty$, whereas the convergence of the estimates of the short-run adjustment parameters is proportional to $\sqrt{T}$ (Johansen, 1996). This gives the justification for performing identification in two steps: (1) the identification of the long-run parameters, β, and (2) the identification of the short-run structure conditional on the identified β. The short-run parameters in (10) and (9) are related by the following equalities:

$$A_1 = A_0 \Gamma_1, a_1 = A_0 \alpha, \mu_{0,a} = A_0 \mu_0, v_t = A_0 \varepsilon_t, \Sigma = A_0 \Omega A_0'.$$

Identification of the long-run structure of r (simultaneous) relations requires at least $r(r-1)$ restrictions, whereas the identification of the short-run simultaneous equations requires at least $p(p-1)$ restrictions. In both cases the restrictions have to satisfy generic identification as measured by the identification rank conditions (Johansen, 1995, Johansen and Juselius, 1994). The identifying restrictions of the long-run structure are formulated as:

$$\beta = \{H_1 \phi_1, ..., H_r \phi_r\} \tag{11}$$

where H_i is a design matrix $p \times m_i$ that satisfies generic identification and ϕ_i is a $m_i \times 1$ vector of unrestricted coefficients. Note that generic identification of (11) does not involve restrictions on the long-run covariance matrix. This is because stationarity now replaces independence as a condition for uniqueness

Identification of the short-run adjustment structure is about how to impose identifying restrictions on the contemporaneous matrix A_0, given lagged changes of the process, Δx_{t-1}, and lagged equilibrium errors, $\hat{\beta}' x_{t-1}$, where $\hat{\beta}$ is an estimated long-run structure. Identifying restrictions on the equations system, $A'X_t = v_t$, where $A' = (A_0, -A_1, -a)$ and $X_t' = (\Delta x_t', \Delta x_{t-1}', x_{t-1}'\beta)$, are formulated by the design matrices H_i, $i = 1, ..., p$:

$$A = (H_1\varphi_1, ..., H_p\varphi_p).$$

Since the short-run adjustment dynamics, $A'X_t$, define relationships between stationary variables, economic identification generally requires the residuals to be uncorrelated. Large off-diagonal elements of the covariance matrix Ω arise when the system variables are simultaneously correlated as a result of a simultaneous relationship between endogenous variables, or of omitted variables simultaneously influencing the system variables.

The next question to be addressed is how identification of the long-run and short-run structure of the CVAR relates to identification in the traditional simultaneous equations model. But, first we need to discuss the concept of irreducibility of a cointegration relation.

4.2 Irreducible cointegration relations and coflux equations

An identified cointegration structure consists of r irreducible cointegration relations (Davidson, 1999) where irreducibility means that stationarity will be lost if one of the variables is omitted from the relation. Thus, an irreducible cointegration relation contains exactly the right number of variables needed to make it stationary, no less, no more. This seems closely related to what Frisch called a coflux equation. John Aldrich explained the latter as:

> An equation which is irreducible with respect to the set of functions which forms the actual solution of the complete system is called a "coflux" equation. In the 1936 language, the coefficients of such an equation have a "statistically uniquely determined meaning".
>
> Coflux equations were significant because they could be estimated (in the deterministic case, solved for) from the data - "they were discoverable through passive observations." In modern terms, coflux equations are the ones that are identified. ...
>
> Coflux equations are estimable but they are not necessarily the most interesting equations. They may only be confluent relations. [Aldrich,1989, p. 24]

To understand how the concept of irreducibility can be associated with autonomy, two additional aspects need to be discussed. First, because a structural relation can be a combination of two (or several) irreducible cointegration relations, economic identification often involves the α coefficients which tie the two relations together. For example, consider the deviations from a money demand relation, $(m-p-y)-b_1(R_s-R_l)$, where m is money stock, p is prices, y is income, and R_s-R_l is the spread between the short and long-term interest rate measuring the opportunity cost of holding money relative to bonds. If $(m-p-y)$ and (R_s-R_l) are both nonstationary and cointegrated, a cointegration relation between the two can directly identify the money demand relation. But if the interest rate spread (R_s-R_l) is stationary by itself (as standard theory would predict) then money velocity $(m-p-y)$ would also have to be stationary for an empirical money demand relation to be stationary. In the latter case the economic relation of interest is a combination of two stationary cointegration relations linked by the α_m coefficients in the money stock equation.

$$\alpha_{m,1}(m-p-y)+\alpha_{m,2}(R_s-R_l) \tag{12}$$

If $\alpha_{m,1}<0$, implying error correcting behavior, and $\alpha_{m,2}<0$, implying that money demand increase when the opportunity cost of holding money goes down, then we may argue that we have identified a money demand relation.

Second, because a linear combination of the r cointegration relations is also a stationary relation, one can find infinitely many 'new' stationary relations by combining them. For example, if $x_{1,t}$ and $x_{2,t}$ as well as $x_{2,t}$ and $x_{3,t}$ are cointegrated, then $x_{1,t}$ and $x_{3,t}$ are also cointegrated. This means that there are often many ways of identifying a long-run structure and not all of them would satisfy economic identification. In this sense, an irreducible cointegration relation seems to share many of the characteristics of a confluent relation.

4.3 Autonomy, structural invariance, and cointegration: a discussion

> The question of what connection there is between the relations we work with in theory and those we get by fitting curves to actual statistical data is a very delicate one. I think it has never been exhaustively and satisfactorily discussed. [Frisch, 1938]

Economic identification requires that the restricted cointegration relations describe theoretically interpretable simultaneous relationships between the VAR variables. However, cointegration only implies that the variables are associated over the long run, whereas it does not, as such, say anything about the direction of causality. Also the VAR does not distinguish between endogenous and exogenous variables: all stochastic variables are modelled and exogeneity is tested within the VAR, not assumed from the outset. This is in contrast to a traditional simultaneous equations system where the variables are either assumed to be endogenous or exogenous and only the endogenous variables are modelled. Causality is assumed from the outset by normalizing each equation on one of the endogenous variables and expressing it as a function of the other endogenous, exogenous, and predetermined variables.

How do we solve the problem of normalization and causal links between the variables in the context of a cointegration relation? While the β relations as such are not informative about causal links between the variables, the α coefficients are. This is because they provide information about how the system adjusts when exogenous shocks have pushed the long-run relations out of equilibrium. The sign of an adjustment coefficient α_{ij} determines whether the variable $x_{i,t}$ in $\beta' x_t$ is error correcting, error increasing, or not adjusting at all in the equation $\Delta x_{i,t}$. Normalizing on a variable $x_{i,t}$ in $\beta'_j x_t$ generally presumes a significant adjustment coefficient α_{ij} preferably describing equilibrium error correction behavior. If the α_{ij} is insignificant, then the variable $x_{i,t}$ can be considered "exogenous" in that relation. Two, or more, variables in $\beta'_j x_t$ exhibiting significant adjustment in their respective equations implies simultaneous feedback. Thus, the economic identification of the β relations (11) is incomplete in terms of causal links unless it is tied to the short-run adjustment structure.

Haavelmo and Frisch argued that a requirement for a relation to be autonomous is that its coefficients remain unchanged when other parts of structure are changing.[6] Is an appropriately identified cointegration relation autonomous in the above sense? It may or may not be. For example, the stationary money demand relation (12) was shown to have remained unchanged from 1975 to 2003 except for a shift in the equilibrium mean in connection with a major financial deregulation in 1983. For details see Juselius (2006). Would Haavelmo/Frisch have considered it an autonomous relation despite the mean shift? Is the relation likely to remain stable under other future

[6] This is closely related to the concept of super exogeneity in Engle et al. (1983).

reforms that change other parts of the economic structure? Such questions call for modesty: cointegration is powerful measure of association when data are non-stationary, but as such it is basically a statistical regularity that may break down if conditions change. It is not difficult to share Frisch conclusion that autonomy and structural invariance are theoretical concepts which are empirically elusive.

Another problem that needs to be addressed is that it is not always possible to impose economically meaningful identifying restrictions on *all* cointegration relations. This is often the case when only some of the variables (the 'endogenous' variables) are theoretically well determined by the chosen information set, whereas the other (the 'exogenous' variables) are only partially determined. If the variables being exogenous in theory would be so also in real life, the problem would not arise. But this is far from always the case and the identified long-run structure is often a mixture of fully determined economic relations and partially specified relations with no structural interpretation.

This might seem as a serious drawback for the cointegration analysis but need not be so. Juselius (2010) showed that by formulating a theory consistent CVAR scenario one can derive a *theoretically consistent* long-run structure that separates *a priori* between fully specified structural relations and partially specified relations with less clear interpretation. In the next section I shall argue that such a CVAR scenario may come close to what Haavelmo had in mind when he argued for the necessity to accompany a theory model with a design of experiment.

5 An design of experiment for passive observations and the CVAR

It was a central idea in Haavelmo's work that a theory model had to be accompanied by a design of experiment that *describes and explains* how to measure the variables supposed to be the true variables:

> A theoretical model in this sense is, as it stands, void of any practical meaning or interest. And this situation is, as we have previously explained, not changed by merely introducing "economic names" for the variable quantities or objects involved. The

> model attains economic meaning only after a corresponding system of quantities or objects in real economic life has been chosen or described, in order to be identified with those in the model. That is, the model will have an economic meaning only when associated with a design of actual experiments that describes - and indicates how to measure - a system of "true" variables (or objects) xl, x2, , xn that are to be identified with the corresponding variables in the theory. (Haavelmo, 1944, p. 8)

The formulation of a theory consistent CVAR scenario (Juselius, 2006, 2010 and Juselius and Franchi, 2007) is an attempt to meet Haavelmo's requirement of a design of experiment for passive observations. The basic ideas will be discussed and illustrated with a monetary model for inflation in Romer (1996, Chapter 9).

5.1 Inflation and excess liquidity

Assume that the purpose of the Cointegrated VAR analysis is to identify and estimate an aggregate money-demand relation in order to assess the effect of excess liquidity on inflation. Monetarist theories predict that the inflation rate is directly related to expansions in the (appropriately defined) supply of money at a rate greater than that warranted by the growth of the real productive potential of the economy. The policy implication is that the aggregate supply of money should be controlled in order to control the inflation rate. The optimal control of money, however, requires knowledge of the "non- inflationary level" of aggregate demand for money at each point of time, defined as the level of money stock, m^*, at which there is no tendency for the inflation rate to increase or decrease. On a practical level, the reasoning is based on the assumption that there exists a stable (autonomous) aggregate demand-for-money relation, $m^* = f(x)$, that can be estimated.

Theory tells us there are three distinct motives for holding money: the transactions motive (the need to hold cash for handling everyday transactions); the precautionary motive (the need to hold money to meet unforeseen expenditures); and the speculative motive (agents' wish to hold money as part of their portfolio). Since all three motives are likely to affect agents' needs to hold money, the initial assumption is that money demand, m, is a function of the level of income, y, (assumed to primarily determine the volume of transactions and precautionary money) and the cost of holding money, $R_l - R_s$ and $\Delta p - R_s$, (assumed to determine the cost of transactions, precautionary and speculative demand) where R_l is a long-term interest rate, R_s is the interest rate yield on money, and Δp is the inflation rate.

Which measurements do we choose for the empirical study and do they correspond to the 'true' variables of the theory? Liquidity is measured by M3 which is a broad monetary aggregate, income by gross national expenditure, GNE, inflation by the change in CPI prices, the short rate by the bank deposit rate and the long rate by the government bond rate. All variables are in logarithmic form except for interest rates and all variables are collected by passive observation. As argued in more detail in Juselius (1993) none of them are likely to closely represent the true variables of the underlying theory. They are primarily chosen because they are available in existing data bases and because they are the variables monetary authorities are supposed to react on. With Haavelmo's words:

> The economist, on the other hand, often has to be satisfied with rough and biased measurements. It is often his task to dig out the measurements he needs from data that were collected for some other purpose; or, he is presented with some results which, so to speak, Nature has produced in all their complexity, his task being to build models that explain what has been observed. [Haavelmo, 1944, p.]

5.2 A stochastic mechanism

> A sample of observations is just a set of cold, uninteresting numbers unless we have a theory concerning the stochastic mechanism that has produced them." [Haavelmo, 1950: p. 265].

The inverted VAR model is a useful representation of the stochastic mechanism that have generated the data. It describes the variables as a function of the shocks to the system and the deterministic variables, such as a constant, time trends, step dummies (and possibly exogenous variables). In modern language it is called the common stochastic trends model, in the language of Frisch and Haavelmo it would have been called the final equations.

When inverting the VAR in (8) the deterministic component ΦD_t is, for simplicity, assumed only to contain an (unrestricted) constant, μ_0:

$$x_t = C\sum_{i=1}^{t}\varepsilon_i + C\mu_0 t + C^*(L)\varepsilon_t + \widetilde{X}_0, \tag{13}$$

where $C = \beta_\perp(\alpha'_\perp(I-\Gamma_1-...-\Gamma_{k-1})\beta_\perp)^{-1}\alpha'_\perp,$ measures the long-run impact of a shock to the system, $C^*(L)$ is an infinite lag polynomial describing the impulse response function, and $\widetilde{X}_0$ contains the initial values, $x_0, x_{-1}, ..., x_{-k+1}$, of the process and the initial value of the short-run dynamics $C^*(L)\varepsilon_0$. The representation (13) describes a decomposition of the vector process, x_t, into stochastic trends, $C\sum_{\mathbf{i}=1}^{t}\varepsilon_i$, deterministic trends, $C\mu_0 t$, cycles, $C^*(L)\varepsilon_t$, and irregular components, ε_t. Such a decomposition into trends and cycles was central to much of Frisch' empirical work about business cycles and their effect on the macroeconomy with the exception that the trends he considered were deterministic rather than stochastic.

6 A theory consistent CVAR scenario

The first step in the formulation of a theory consistent CVAR scenario is to specify the theory-consistent number of autonomous stochastic and deterministic trends in (13) and how they load onto the selected variables. As argued in Juselius (2006), the "Rational Expectations" monetary model of inflation in Romer is *implicitly* assumed to be driven by two structural shocks u_1 and u_2, where for simplicity u_1 can be assumed a nominal shock causing a permanent shift in the aggregate demand curve and u_2 is a real shock causing a permanent shift in the aggregate supply curve.

Next, we need to formulate hypotheses for how the stochastic mechanism is supposed to unfold under the given theory. This leads to a first tentative decomposition of the data vector into two stochastic trends, one deterministic time trend, and a stationary cycle component:

$$\begin{bmatrix} m_t \\ p_t \\ y_t^r \\ R_{m,t} \\ R_{b,t} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ 0 \\ 0 \\ 0 \end{bmatrix} [\sum\sum u_{1i}] + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \\ d_{41} & d_{42} \\ d_{51} & d_{52} \end{bmatrix} \begin{bmatrix} \sum u_{1i} \\ \sum u_{2i} \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ 0 \\ 0 \end{bmatrix} [t] + \text{stat.comp.} \tag{14}$$

Under the assumption that inflation, Δp_t, is integrated of order one (consistent with empirical results) the nominal shock $u_{1,t}$ cumulates twice to describe a second order stochastic trend in money and prices and once to describe a first order stochastic inflation trend. The real shock $u_{2,t}$ is assumed to cumulate once consistent with real variables being integrated of order one. The deterministic trend component describes long-term linear growth. The assumption that $\{g_1 \neq 0, g_2 \neq 0, g_3 \neq 0\}$ is consistent with the fact that nominal money, prices, and real income have exhibited linear growth over the sample period.

Long-run price homogeneity, an important feature of monetary models, is formulated as the restriction $c_{11} = c_{21}$. Under this assumption, the nominal-to-real transformation (Kongsted, 2005) applies allowing us to reformulate (14) in its (almost) equivalent form:

$$\begin{bmatrix} m_t - p_t \\ \Delta p_t \\ y_t^r \\ R_{m,t} \\ R_{b,t} \end{bmatrix} = \begin{bmatrix} d_{11} - d_{21} & d_{12} - d_{22} \\ c_{21} & 0 \\ d_{31} & d_{32} \\ d_{41} & d_{42} \\ d_{51} & d_{52} \end{bmatrix} \begin{bmatrix} \sum u_{1i} \\ \sum u_{2i} \end{bmatrix} + \begin{bmatrix} g_1 - g_2 \\ 0 \\ g_3 \\ 0 \\ 0 \end{bmatrix} [t] + ... \tag{15}$$

All variables in (15) are now at most $I(1)$.

6.1 A CVAR scenario for the long-run structure

The next step is to impose further theory consistent restrictions on (15). As argued in Juselius (2006), the monetary model for inflation is implicitly based on the expectations hypothesis and the Fisher parity. Instead of requiring that the parities hold exactly we assume that they hold as stationary conditions. A stationary term structure implies that the interest rates share one common stochastic trend and, hence, that the interest rate spread, $R_m - R_b$, measuring the opportunity cost of holding money relative to bonds, is stationary. A stationary Fisher hypothesis implies that the real interest rate, $R_m - \Delta p$, measuring the opportunity cost of holding money relative to inflation, is stationary. Thus, money velocity $m - p - y^r$ must also be stationary for the money demand relation to hold as a stationary relation. Thus, a theory consistent CVAR scenario for the long-run structure can be formulated as:

$$\begin{bmatrix} m_t - p_t \\ \Delta p_t \\ y_t^r \\ R_{m,t} \\ R_{b,t} \end{bmatrix} = \begin{bmatrix} 0 & d_{12} \\ c_{21} & 0 \\ 0 & d_{12} \\ c_{21} & 0 \\ c_{21} & 0 \end{bmatrix} \begin{bmatrix} \sum u_{1i} \\ \sum u_{2i} \end{bmatrix} + \begin{bmatrix} g \\ 0 \\ g \\ 0 \\ 0 \end{bmatrix} [t] + ... \qquad (16)$$

The three irreducible cointegration relations can be represented by

$$\begin{aligned} m - p - y &\sim I(0), \\ R_{m,t} - R_{b,t} &\sim I(0), \\ R_{m,t} - \Delta p_t &\sim I(0), \end{aligned}$$

describing a theory consistent world where real income and real money stock share the real trend, $\sum u_{2i}$, and the deterministic time trend and inflation and the two nominal interest rates share the nominal trend, $\sum u_{1i}$.

6.2 A CVAR scenario for the short-run structure

The final step in the scenario describes how the system responds when exogenous shocks have pushed the identified long-run relations away from their equilibrium values. While theoretical models are often informative about restrictions on the long-run structure, they are more silent about the short-run adjustment dynamics. But, even though this is also the case with our example, most monetary models would be consistent with the following hypothetical dynamic adjustment structure:

$$A_0 \begin{bmatrix} \Delta m_t^r \\ \Delta y_t^r \\ \Delta^2 p_t \\ \Delta R_m \\ \Delta R_{b,t} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \\ a_{31} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & a_{52} & 0 \end{bmatrix} \begin{bmatrix} (m - p - y^r)_{t-1} - \beta_{0,1} \\ (R_m - R_b)_{t-1} - \beta_{0,2} \\ (R_m - \Delta p)_{t-1} - \beta_{0,3} \end{bmatrix} + \begin{bmatrix} \gamma \\ \gamma \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \end{bmatrix}. \qquad (17)$$

It is based on the assumption that real income and the short-term interest rate can be considered weakly exogenous for this information set. Moreover, if $a_{11} < 0, a_{12} > 0, a_{13} > 0$, then $(m-p-y^r)+a_{12}/a_{11}(R_m-R_b)+a_{13}/a_{11}(R_m-\Delta p)$ can be economically identified as a money demand relation; if $a_{31} > 0$ then inflation will increase when there is excess liquidity in the economy; if $a_{41} < 0$, then the bank deposit rate will decrease when excess liquidity increases, and if $a_{43} < 0$ it will increase when inflation increases; and if $a_{52} > 0$, then the long-term bond rate will adjust to the short-long spread as the expectations hypothesis mostly assume. The common stochastic trends are given by the cumulated shocks to real income, $\sum \varepsilon_{y^r}$, and to the short-term interest rate, $\sum \varepsilon_{R_m}$.

6.3 A design of experiment for passive observations: a discussion

I have argued above that reliable inference on the effect of excess liquidity on inflation and the rest of the economy requires a 'design of experiment' for the whole system formulated as a long-run structure of irreducible cointegration relations combined with a short-run structure of adjustment dynamics that indicates how the system is expected to adjust when out of equilibrium. This in my view is the closest we can come to Haavelmo's recipe for testing a theory when data are obtained by passive observation:

> The idea behind this is, one could say, that Nature has a way of selecting joint value-systems of the "true" variables such that these systems are as if the selection had been made by the rule defining our theoretical model. Hypotheses in the above sense are thus the joint implications and the only testable implications, as far as observations are concerned, of a theory and a design of experiments. It is then natural to adopt the convention that a theory is called true or false according as the hypotheses implied are true or false, when tested against the data chosen as the "true" variables. Then we may speak, interchangeably, about testing hypotheses or testing theories. [Haavelmo, 1944, p. 9]

Though appealing from a theory point of view, (16) and (17) did not received much empirical support in (Juselius, 2006). Instead, real interest rates, interest rate spreads, and money velocity were found too persistent to support stationarity and the long-term interest rate, rather than the short term, was found weakly exogenous similar as in Giese (2008). Therefore, to become empirically relevant, the monetary model in Romer (1996) would have to account for these unexplained properties of the data. For example, equilibrium mean shifts in the historical data such as the structural shift in 1983 means that economic actors had not been able to foresee the consequences of the reform and, therefore, had not been able to act in a way that would have mitigated its effect on inflation, income, and money stock, say. If economic actors were not able to foresee the effect of such changes in the past, it seems rather unlikely they will do so in the future. Thus, unconditional forecasts based on a model that is subject to a future change in its distribution (such as a location shift) is likely to suffer from systematic bias. Such features of the data are problematic for Rational Expectations Hypothesis (REH) based models as they postulate a universally constant parameter model based on which economic actors recursively form expectations. Romer's (1996) model relies on this hypothesis.

These features of the data (pronounced persistence and breaks) are more in line with the recent theory of Imperfect Knowledge Economics (IKE) (Frydman and Goldberg, 2007, 2012). The REH models preassume that nonroutine change, i.e. change that cannot be specified in advance with mechanical rules and procedures, is unimportant for understanding participants' decision making and market outcomes (Frydman and Goldberg, 2012), whereas nonroutine change is central for the IKE theory. Given imperfect knowledge, distributions are likely to change as the world changes but in a way that cannot be specified in advance. Thus, even though a distribution approximates the observables over a given sample period IKE supposes, as Keynes and Frisch did, that there is no reason to believe it holds universally in the past, present and the future.

The CVAR assumption of multivariate normality is designed to hold *ex post* for "the sample as it is" and often, as in the present case, only after having controlled for the effect of extraordinary events that have shifted the mean of the distribution. Also, the persistent deviations from long-run parity conditions that remained unexplained based on the REH monetary model of Romer (1996) can be rationalized by the IKE theory. Frydman and Goldberg (2007) show that, under imperfect knowledge, economic actors tend to push prices persistently away from long-run benchmark values, thereby creating such pronounced persistence as we frequently see in the data.

The above discussion has served the purpose of illustrating that a theory consistent CVAR scenario allows us to assess the empirical validity of the basic assumptions underlying a chosen theory model (formulated as testable assumptions on the pulling and pushing forces) and, hence, how and where to modify the theory. The latter was an integral part of Haavelmo's methodological thinking:

> In order to test a theory against facts, or to use it for predictions, either the statistical observations available have to be "corrected," or the theory itself has to be adjusted, so as to make the facts we consider the "true" variables relevant to the theory, as described above. [Haavelmo, 1944 p. 7]

7 A concluding discussion

This paper has tried to demonstrate that many econometric problems which were discussed by Haavelmo and his contemporaries such as time dependent residuals, spurious correlation and regression, multicollinearity, normalization, reduced rank, and to some degree, missing variables can be given a practical solution within the general framework of a *well* specified CVAR model. The latter was derived by specifying the joint probability for all observables, i.e. by following Haavelmo's overriding idea for how to do econometric modelling.

The paper demonstrates how the nonstationarity of economic data allows us to formulate identification, simultaneity, and structural invariance in a much richer context than was possible in Haavelmo's time. But because "data by passive observation" do not generally qualify for the purpose of testing "deep structural parameters" of theoretical models the paper concludes that modesty is needed in particular with respect to structural invariance.

The frequent finding that the normality assumption and parameter constancy are acceptable only *ex post* after we have allowed for shifts in the equilibrium mean (or shifts in the growth rates) due to extraordinary institutional events implies that the CVAR is not likely to produce unbiased forecast errors over periods potentially subject to structural changes and location shifts in the probability distribution. The paper argues that this may *a priori* have serious implications regarding the class of theoretical models that potentially can be considered empirically relevant.

Haavelmo's argued that a theoretical model should always be accompanied with a design of experiment explaining how to adequately take the theory to data obtained by passive observations. The paper proposes a theory-consistent CVAR scenario as a design of experiment for passive observations arguing that it translates as many as possible of the assumptions of a theoretical model into testable hypotheses on the 'pulling and pushing' forces of the data generating process. Such a scenario may also provide a common modeling strategy that allows macroeconomic questions to be investigated in a consistent framework. For example, one might be able to learn from the "experiments" provided by other countries that differ in various aspects with regard to the investigated economic problem. Most importantly, a theory consistent CVAR scenario is often informative about how to modify the theory model when the correspondence between the theoretical and observed structure is weak. In Haavelmo's words:

> ... In the second case we can only try to adjust our theories to reality as it appears before us. And what is the meaning of a design of experiment in this case. It is this: We try to choose a theory and a design of experiments to go with it, in such a way that the resulting data would be those which we get by passive observation of reality. And to the extent that we succeed in doing so, we become masters of reality - by passive agreement. [Haavelmo, 1944, p.14]

8 References

Aldrich, J. (1989), "Autonomy", *Oxford Economic Papers*, New Series, Vol. 41, No. 1, History and Methodology of Econometrics, pp. 15-34.

Anderson, T.W. (1951), "Estimating linear restrictions on regression coefficients for multivariate normal distributions", *Annals of Mathematical Statistics,* 22, 327-351.

Andersen, T.W. (1991). Tryggve Haavelmo and Simultaneous Equation Models. *Scandinavian Journal of Statistics*, 18, 1-19.

Bjerkholt, O. (2011), "Trygve Haavelmo, Frisch and the probability approach", the Department of Economics, University of Oslo.

Davidson, J. (1988), "Structural relations, cointegration and identification: Some simple results and their application", *Journal of Econometrics*, 87(1), 87-113.

Engle, R.F., D.F. Hendry, and J-F Richard (1983), "Exogeneity", *Econometrica*, 51, 2, 277-304.

Frish, R. (1938), "Statistical versus theoretical relations in economic macrodynamics", Unpublished memorandum, University of Oslo.

Frydman, R. and M. Goldberg (2007), *Imperfect Knowledge Economics: Exchange rates and Risk, Princeton.* NJ: Princeton University Press.

Frydman, R. and M. Goldberg (2011), *Beyond Mechanical Markets: Risk and the Role of Asset Price Swings*, Princeton University Press, Princeton.

Garrone, G., R. Marchionatti, and R. Bellofiore (2004). "Keynes on Econometric Method. A Reassessment of his Debate with Tinbergen and Econometricians, 1938-1943," CESMEP Working Papers 200401, University of Turin

Giese, J.V. (2008), "Level, Slope, Curvature: Characterizing the Yield Curve in a Cointegrated VAR Model". *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 2, 28, http://www.economics-ejournal.org/economics/journalarticles/2 28.

Haavelmo, T. (1943), "The statistical implications of a system of simultaneous equations", *Econometrica*, 11, 1-12.

Haavelmo, T. (1944), "The Probability Approach to Econometrics". *Econometrica*, 12 (Supplement), 1-118.

Haavelmo, T. (1950), "Remarks on Frisch's Confluence Analysis and its Use in Econometrics", in *Statistical Inference in Dynamic Models,* ed. T. C. Koopmans, Cowles Commission Monograph, Vol. 10, John Wiley & Sons, New York, 258-265.

Hansen, H. and S. Johansen (1999). "Some tests for parameter constancy in cointegrated VAR-models". *The Econometrics Journal*, 2(2), 306–333.

Hendry, D.F. and G.E. Mizon (1993), "Evaluating Econometric Models by Encompassing the VAR" in Phillips, P.C. (ed.) *Models, Methods and Applications of Econometrics*, Blackwell, Oxford, 272-300.

Hendry, D.F. and M.S. Morgan (1989), "A Re-Analysis of Confluence Analysis", *Oxford Economic Papers*, Vol. 41, No. 1, History and Methodology of Econometrics, pp. 35-52

Hendry, D.F. and M.S. Morgan (1995), "*The Foundations of Econometric Analysis*", Cambridge University Press, Cambridge.

Hendry, D.F. and J.F. Richard (1983), "The econometric analysis of economic time-series (with discussion), *International Statistical Review*, 51, 111-163.

Hendry, D.F., Spanos, A. and Ericsson, N.R. (1989). The Contributions to Econometrics in Tryggve Haavelmo's 'The Probability Approach in Econometrics'. *Socialøkonomen*, 43(11), 12-17.

Johansen, S. (1988). "Statistical Analysis of Cointegration Vectors". *Journal of Economic Dynamics and Control*, 12(213), 231–254.

Johansen, S. (1995), "Identifying restrictions of linear equations. With applications to simultaneous equations and cointegration. *Journal of Econo-*

metrics, 69(1), 111-132.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Oxford*. Oxford University Press.

Johansen, S. and K. Juselius (1994), "Identification of the long-run and short-run structure: An application to the ISLM model", *Journal of Econometrics*, 63, 7-36.

Juselius, K. (1993), "VAR Modelling and Haavelmo's Probability Approach to Macroeconomic Modelling", *Empirical Economics*,18, 595-622.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*. Oxford: Oxford University Press.

Juselius, K. (2010), "Testing Exchange Rate Models Based on Rational Expectations versus Imperfect Knowledge Economics: A Scenario Analysis". Working Paper, Economics Department, University of Copenhagen.

Juselius, K. and M. Franchi (2007), "Taking a DSGE Model to the Data Meaningfully." *Economics-The Open-Access, Open-Assessment E-Journal*, No. 4, http://www.economics-ejournal.org/economics/journalarticles/2007-4.

Kongsted, H. C. (2005), "Testing the nominal-to- real transformation", *Journal of Econometrics*, 124, 202-225.

Nielsen, H. B. (2008), "Influential Observations in Cointegrated VAR Models: Danish Money Demand 1973-2003". *The Econometrics Journal*, 11 (1): 1-19.

Phillips, P.C.B., (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics* 33, 311-340.

Romer, D. (1996). *Advanced Macroeconomics.* McGraw Hill: New York.

Yule, U. (1926), "Why do we Sometimes Get Nonsense-Correlations Between Time Series? – A Study in Sampling and the Nature of Time Series". *Journal of the Royal Statistical Society* 89, 1–63.