DEPARTMENT OF ECONOMICS

Working Paper

MULTICOINTEGRATION IN
STOCK-FLOW MODELS

Tom Engsted
Niels Haldrup

Working Paper No. 1997-18
Centre for Non-linear Modelling in Economics

ISSN 1396-2426

UNIVERSITY OF AARHUS • DENMARK

CENTRE FOR NON-LINEAR MODELLING IN ECONOMICS
DEPARTMENT OF ECONOMICS - UNIVERSITY OF AARHUS - DK - 8000 AARHUS C - DENMARK
☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

WORKING PAPER

MULTICOINTEGRATION IN STOCK-FLOW MODELS

Tom Engsted
Niels Haldrup

Working Paper No. 1997-18

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND MANAGEMENT - UNIVERSITY OF AARHUS - BUILDING 350
8000 AARHUS C - DENMARK ☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Multicointegration in Stock-Flow Models

Tom Engsted*
Department of Finance
Aarhus School of Business, DK-8210 Aarhus V, Denmark
Niels Haldrup
Department of Economics
University of Aarhus, DK-8000 Aarhus C, Denmark

13 October, 1997

Abstract. Multicointegration, in the sense of Granger and Lee (1990), frequently occurs in models of stock-flow adjustment and implies cointegration amongst I(2) variables and their differences (polynomial cointegration). The purpose of this article is two-fold. First, we demonstrate that based on a multicointegrated vector autoregression (VAR) two equivalent error correction model (ECM) representations can be derived; the first is expressed in terms of adjustments in the flows of the variables (the standard I(2) ECM), and the second is expressed in terms of adjustments in both the stocks and the flows. Secondly, we apply I(2) estimation and testing procedures for multicointegrated time series to analyze data for US housing construction. We find that stocks of housing units started and completed exhibit polynomial cointegration (and hence the flows are multicointegrated) and the associated ECM's are estimated. Lee (1992, 1996) also found multicointegration in this data set but without explicitly exploiting the I(2) property.

Keywords: Cointegration, multicointegration, I(2) processes, housing construction.
JEL Codes: C12, C13, C22, C32, C51.

1. Introduction

Granger and Lee (1989, 1990) have defined the notion of multicointegration amongst economic variables integrated of order one: If the cumulation of cointegration (equilibrium) errors cointegrate with the original variables the series are said to be multicointegrated. The phenomenon is most likely to occur in models describing the

*Both authors are affiliated with Centre for Non-Linear Modelling in Economics at the University of Aarhus. Part of this work was carried out while the second author was visiting the European University Institute in Florence during the spring, 1997. We appreciate helpful comments from Søren Johansen and we thank Tae-Hwy Lee for providing us with the US housing construction data set.

dynamic interactions between stock and flow variables. Multicointegration implies that in the bivariate I(1) model, for example, there can exist more than just a single cointegration vector. More generally, the number of cointegrating relations and the number of common stochastic trends will not add up to the dimension of the system which is the case in the usual cointegrated I(1) model. In a recent paper Engsted and Johansen (1997) show that when variables are multicointegrated the requirements for the system to be an I(1) system will fail; in fact, an I(1) specification will be misspecified even though the main interest lies in the analysis of the I(1) series. Instead the system should be formulated as an I(2) model where multicointegration can be shown to result in cointegration amongst generated I(2) variables and their first differences (polynomial cointegration)[1]. Others have been aware of the link between multicointegration and polynomial cointegration, see *inter alia* Engle and Yoo (1991), and Lee (1992), but to the best of our knowledge there have been no empirical contributions to the multicointegration literature which takes explicitly account of the I(2) property in estimation and hypothesis testing. In this paper we provide an example of how such an analysis can be undertaken.

We also want to show how multicointegrated variables from a general vector autoregression (VAR) can be given two equivalent error correction model (ECM) representations. The first representation is a direct consequence of the I(2) formulation of the model and follows Johansen's (1992) version of Granger's representation theorem for I(2) variables. It describes how the flow variables adjust in response to disequilibrium errors between the stocks and the flows as well as between the flow variables themselves. The second ECM representation describes adjustments in *both* the stock *and* the flow variables in response to the equilibrium errors. Although equivalent, the two distinct ways of writing the ECM provide different insights of the complex dynamics characterizing multicointegrated systems.

The plan of the paper is the following. In section 2 we briefly summarize some known results for the relationship between optimum control models and error correction models. We also demonstrate how multicointegrating relationships can be given an I(2) formulation. The next section describes the relationship between multicointegration and cointegrated I(2) VAR models and the section also serves to define the notation to be used subsequently in section 4 where we show how the two different ECM representations for multicointegrated time series can be obtained. In section 5 we demonstrate the I(2) analysis on the US housing construction data set 1968:1-1994:12 which has previously been examined by Lee (1996) (and Lee (1992) for a

[1] In the literature the notions of 'multicointegration' and 'polynomial cointegration' are frequently used as synonyms. However, we use 'multicointegration' in the particular sense of Granger and Lee, and hence this notion applies to systems where the basic variables are really I(1) but can be transformed to an I(2) system. On the other hand, 'polynomial cointegration' refers to cointegrated models where the basic variables are I(2) and the levels and the first differenced variables cointegrate.

shorter data period) but without exploiting the I(2) property. This data set is found to nicely characterize a multicointegrating relationship amongst the variables and we calculate the associated ECMs. The final section concludes.

2. Optimum control ECM's and the definition of stock variables

Granger and Lee (1990) have shown how the ECM for multicointegrated time series can be derived as the solution to a particular control problem[2]. In a slightly more general context Lee (1996) addresses the infinite horizon linear quadratic adjustment cost model

$$L_t = \mathsf{E}_t \sum_{j=0}^{\infty} \delta^j [(x_{t+j} - y_{t+j})^2 + \lambda_1 (Q_{t+j} - \kappa y_{t+j})^2 + \lambda_2 (x_{t+j} - x_{t+j-1})^2] \tag{1}$$

where, assuming a zero initial condition, $Q_t = Q_{t-1} + (x_t - y_t) = \sum_{j=0}^{t} (x_j - y_j)$. The variable x_t is frequently referred to as the control variable whereas y_t is the target series one is attempting to track by choice of x_t. The basic I(1) flow variables are x_t and y_t while Q_t is the I(1) stock variable which measures the cumulated discrepancies between the control and the target series. δ is the discount rate and λ_1, λ_2 are cost parameters. Note that $x_t - y_t$ and $Q_t - \kappa y_t$ are cointegrating relations and hence are I(0). Lee (1996) shows how the minimization of the loss function (1) leads to a stock adjustment model which essentially is of the form:

$$\Delta Q_t = \alpha + \beta (Q_{t-1} - \kappa y_{t-1}) + \gamma (x_{t-1} - y_{t-1}) + \mu (y_t - y_{t-1}) \tag{2}$$

where β, γ, and μ are complicated non-linear functions of δ, λ_1, and λ_2. This type of models is familiar from the control engineering literature and belongs to the general class of feedback controllers with proportional, integral, and derivative corrections, see Kwakernaak and Sivan (1972) and Phillips (1954). It also introduces the idea of an inventory accelerator mechanism, see Metzler (1941) and Holt *et al.* (1960). Notice that the ECM (2) is defined in terms of adjustments in the stock variable Q_t.

The adjustment rule derived in Granger and Lee (1990) is based on a one period decision problem and corresponds to an ECM for the flow variables, e.g.

$$\Delta x_t = \widetilde{\alpha} + \widetilde{\beta} (Q_{t-1} - \widetilde{\kappa} y_{t-1}) + \widetilde{\gamma} (x_{t-1} - y_{t-1}) + \widetilde{\mu} (y_{t-1} - y_{t-2}). \tag{3}$$

The decision rules (2) and (3) provide the motivation for the present paper. The idea is that from an optimization point of view it is reasonable to consider adjustment mechanisms for both the stock and the flow variables in this class of models. As such, (2) and (3) motivate the different ECM mechanisms that can be derived within

[2]Hendry and von Ungern-Sternberg (1981) report similar results but without using the notion of cointegration because this concept was not developed until the work of Granger (1983).

a general framework for multicointegrated VAR models. We want to show how the two sorts of representations are indeed equivalent and hence will capture the basic features of the above adjustment mechanisms[3].

There are many examples of stock-flow models that can be considered as particular cases of the above decision rules. Granger and Lee (1989, 1990) address the case where x_t and y_t are production and sales of some commodity, and $\Delta Q_t = x_t - y_t$ is the change of inventories. The Hendry and von Ungern-Sternberg (1981) study of the consumption function belongs to the same class of models.

Consider also the following bivariate example which has been analyzed by Lee (1992) in a BULLETIN article and in Lee (1996) for an extended sample period: We focus on housing market data and let x_t and y_t be the flows of housing units started and completed, respectively. For the current period the increment in the stock of housing units under construction is given by $x_t - y_t$ and the entire stock (assuming a zero initial stock) is $Q_t \equiv \sum_{j=1}^{t}(x_j - y_j)$. Also in this case the potentially cointegrating relations are between the variables x_t and y_t (with an obvious cointegration vector given by $(1, -1)$), and between the variables Q_t and y_t (or possibly x_t). However, the identity defining the increment in the stock of housing units under construction ΔQ_t needs not always be given simply as $x_t - y_t$.

> "It could be [that] some building companies go bankrupt and some starts are therefore not completed, and it could also be due to recording errors. The possibility that dwellings disappear by demolition, fire or conversion to other uses would also affect the identity", (Lee, 1992, p. 423).

The percentage of never completed starts is naturally an unknown figure. Lee solves this problem by using a smoothing estimator and for US data he finds that approximately 2 % of all housing units started are never completed and hence the appropriate stock variable is redefined as $Q_t \equiv \sum_{j=1}^{t}(.98x_j - y_j)$.

The $\mathsf{I}(2)$ representation of the model suggests a different way of estimating this parameter. The idea is to consider the generated $\mathsf{I}(2)$ variables $Y_t = \sum_{j=1}^{t} y_j, X_t = \sum_{j=1}^{t} x_j$, and the $\mathsf{I}(1)$ variables $\Delta Y_t = y_t$, $\Delta X_t = x_t$ (upper case letters represent the cumulated series) and the associated relation

$$Y_t = \kappa_0 X_t + \kappa_1 \Delta Y_t + \kappa_2 \Delta X_t + u_t \tag{4}$$

which is potentially a multicointegrating (polynomially cointegrating) relation[4] when $u_t \sim \mathsf{I}(0)$. The parameter κ_0 is the estimated percentage of starts that are eventually

[3] The decision rules (2) and (3) entail certain parameter restrictions for the underlying optimizing model to be valid. It is beyond the purpose of the present paper to address this particular issue but merely we want to emphasize how the basic control mechanisms will arise in a multicointegrated VAR model.

[4] Observe that in (4) the cumulated variables X_t and Y_t are allowed to cointegrate with both of

completed. Note that when $\kappa_0 = 1$, (4) simplifies to $Q_t = X_t - Y_t = -(\kappa_1 y_t + \kappa_2 x_t + u_t)$. In the present context, given that the model is a satisfactory description of the data, κ_0 can be estimated at the super-super consistent rate, $O_p(T^{-2})$ where T is the sample size and hence, from an estimation point of view, there are very good statistical reasons for formulating the model in terms of an I(2) model. Observe that since the generated I(2) series Y_t, and X_t generally will have a linear trend it may also be necessary to allow for deterministic components in (4).

Another example where the I(2) representation is useful in estimating unknown parameters defining the stock variables is the life cycle hypothesis of consumption. According to this model consumption cointegrates with income and wealth where the latter, by definition, equals accumulated savings, i.e. the accumulated difference between income and consumption. Hence, there is multicointegration between consumption and income. However, in order to find an appropriate measure of savings, it is frequently necessary in this type of models to estimate the relationship between income (decomposed into labour and capital income) and the value of non-durables consumption and service flows, see *e.g.* Campbell (1987, section 4) for details. An $O_p(T^{-2})$ consistent estimate of this relationship can be obtained using the I(2) cointegration analysis. In contrast, Campbell (1987) estimates this relationship in a standard I(1) analysis and hence only obtains an $O_p(T^{-1})$ consistent estimate.

3. MULTICOINTEGRATION AND I(2) COINTEGRATION

The cointegrated VAR model for I(1) models can be written in the well-known vector error correction model (VECM) form, see Johansen (1991),

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{x}_{t-i} + \boldsymbol{\varepsilon}_t \tag{5}$$

where $\mathbf{x}_t$ is a p dimensional vector I(1) time series (here assumed to consist of flow variables) and $\mathbf{\Pi}$ is a matrix determining the cointegration rank of the VAR. $\boldsymbol{\varepsilon}_t$ is a sequence of *i.i.d.* zero mean errors with non singular covariance matrix $\mathbf{\Omega}$. If rank$\mathbf{\Pi} = r$ the $\mathbf{x}_t$ series has r cointegrating relations and $p - r$ common stochastic I(1) trends driving the system. In this case the matrix $\mathbf{\Pi}$ can be decomposed as $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are both $p \times r$ matrices. The cointegrating relations are given by $\boldsymbol{\beta}'\mathbf{x}_t$ and the associated adjustment coefficients summarized in $\boldsymbol{\alpha}$. Engsted and Johansen (1997) show that if $\mathbf{x}_t$ is multicointegrated such that the cumulated equilibrium errors $\sum_{i=1}^{t} \boldsymbol{\beta}'\mathbf{x}_j$ cointegrate with $\mathbf{x}_t$, then the representation (5) is invalid. In fact, the (long-run) covariance matrix defined for the cointegration relations and the differenced

the variables x_t and y_t. Strictly speaking only one of series x_t or y_t is needed for the cointegration property to hold, compare the adjustment models (2) and (3). However, as we shall see it follows as a natural output from the statistical analysis that the specification (4) is useful when the variables are considered from a systems approach using the Johansen ML procedure.

variables will be singular in this case but alternatively the model can be formulated as a VAR for cointegrated I(2) variables. If we, as we did in section 2, let upper case letters denote cumulated (lower case) series then $\mathbf{X}_t \equiv \sum_{j=1}^{t} \mathbf{x}_j$ follows a p dimensional VAR for I(2) variables with the VECM representation

$$\Delta^2 \mathbf{X}_t = \mathbf{\Pi X}_{t-1} - \mathbf{\Gamma} \Delta \mathbf{X}_{t-1} + \sum_{i=1}^{k-2} \mathbf{\Psi}_i \Delta^2 \mathbf{X}_{t-i} + \boldsymbol{\nu}_t \tag{6}$$

where $\mathbf{\Gamma} = \mathbf{I} - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i$ and $\boldsymbol{\nu}_t$ is an error having non-singular covariance matrix. The I(2) VAR is introduced in Johansen (1992, 1995) and he shows that the parameters satisfy

$$\mathbf{\Pi} = \boldsymbol{\alpha\beta}', \quad \text{with } \boldsymbol{\alpha}, \boldsymbol{\beta}\ p \times r, \quad r < p \tag{7}$$

$$\boldsymbol{\alpha}_\perp' \mathbf{\Gamma} \boldsymbol{\beta}_\perp = \boldsymbol{\xi\eta}', \quad \text{with } \boldsymbol{\xi}, \boldsymbol{\eta}\ (p-r) \times s, \quad s < (p-r) \tag{8}$$

where $\alpha_\perp$ denotes the orthogonal complement of α and hence has dimension $p \times (p-r)$ and satisfies $\alpha_\perp' \alpha = 0$. Paruolo (1996) denotes the numbers r, s, and $p-r-s$ the integration indeces of the I(2) VAR model: r relations cointegrate to I(0) level, possibly by including the first differences of $\mathbf{X}_t$, s relations are I(1) and constitute the I(1) common stochastic trends, and finally, $p-r-s$ relations are I(2) which we refer to as the I(2) common trends. Associated with these three cases it is possible to find mutually orthogonal matrices $(\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ where each component individually provides a basis for the I(0), I(1), and I(2) directions, respectively.

The p different relations can also be written as

$$r : \boldsymbol{\beta}' \mathbf{X}_t - \boldsymbol{\delta\beta}_2' \Delta \mathbf{X}_t \sim \mathsf{I}(0) \tag{9}$$

$$s : \boldsymbol{\beta}_1' \mathbf{X}_t \sim \mathsf{I}(1) \tag{10}$$

$$p-r-s : \boldsymbol{\beta}_2' \mathbf{X}_t \sim \mathsf{I}(2). \tag{11}$$

Note that the I(0) relations can be expressed in terms of linear combinations between the levels of $\boldsymbol{\beta}' \mathbf{X}_t$ (which generally are I(1)) and the differenced I(2) trends $\boldsymbol{\beta}_2' \Delta \mathbf{X}_t$ (which are also I(1)), and hence include the polynomially cointegrating relations. The matrix $\boldsymbol{\delta}$ has dimension $r \times (p-r-s)$ and has the orthogonal complement $\boldsymbol{\delta}_\perp$ of dimension $r \times (r-(p-r-s))$ satisfying $\boldsymbol{\delta}_\perp' \boldsymbol{\delta} = \mathbf{0}$. More precisely $\boldsymbol{\delta}$ is defined as $\boldsymbol{\delta} = \overline{\boldsymbol{\alpha}}' \mathbf{\Gamma} \overline{\boldsymbol{\beta}}_2$ where $\overline{\boldsymbol{\alpha}} = \boldsymbol{\alpha}(\boldsymbol{\alpha}'\boldsymbol{\alpha})^{-1}$ and $\overline{\boldsymbol{\beta}}_2$ is defined in a similar way[5].

Not all the I(0) relations in (9) need to be polynomially cointegrating. The linear combinations where the differenced variables are not needed to yield stationarity are given by $\boldsymbol{\delta}_\perp' \boldsymbol{\beta}' \mathbf{X}_t$. On the other hand, the relations which surely require the

[5] Note that $\mathbf{P}_\alpha = \overline{\boldsymbol{\alpha}} \boldsymbol{\alpha}' = \boldsymbol{\alpha} \overline{\boldsymbol{\alpha}}'$ denotes the projection onto the space spanned by the columns of $\boldsymbol{\alpha}$.

differenced variables are the polynomially cointegrating relations and are given by $\boldsymbol{\delta}'\boldsymbol{\beta}'\mathbf{X}_t - \boldsymbol{\delta}'\boldsymbol{\delta}\boldsymbol{\beta}_2'\Delta\mathbf{X}_t$. These are the truely multicointegrating relations. Hence the r $\mathsf{I}(0)$ relations can be partitioned into $p-r-s$ multicointegrating relations and the remaining $r-(p-r-s)$ cointegrating relations. Of course this separation requires that $r > p-r-s$. In practical situations with multicointegration it seems less likely that the directly stationary cointegrating relations $\boldsymbol{\delta}_\perp'\boldsymbol{\beta}'\mathbf{X}_t$ will be present. This would imply that the differenced linear combinations $\boldsymbol{\delta}_\perp'\boldsymbol{\beta}'\mathbf{x}_t$ be $\mathsf{I}(-1)$. Hence we believe that it is most common that $r = p-r-s$ when the focus is on multicointegration in the sense of Granger and Lee. Finally, it should be mentioned that the $\mathsf{I}(1)$ trends $\boldsymbol{\beta}_1'\mathbf{X}_t$ will be stationary in first differences, *i.e.* $\Delta\boldsymbol{\beta}_1'\mathbf{X}_t = \boldsymbol{\beta}_1'\mathbf{x}_t \sim \mathsf{I}(0)$.

Observe that for bivariate systems certain simplifications will arise. In particular, if the series are multicointegrated then $(r, s, p-r-s) = (1, 0, 1)$ and hence no $\mathsf{I}(1)$ trends can be present.

4. Error correction representation of multicointegrated variables

The VECM for general $\mathsf{I}(2)$ systems (6) can be given the following alternative representation due to Paruolo and Rahbek (1996):

$$\Delta^2\mathbf{X}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathbf{X}_{t-1} - \boldsymbol{\delta}\boldsymbol{\beta}_2'\Delta\mathbf{X}_{t-1}) - (\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)(\boldsymbol{\beta}, \boldsymbol{\beta}_1)'\Delta\mathbf{X}_{t-1} + \sum_{i=1}^{k-2}\boldsymbol{\Psi}_i\Delta^2\mathbf{X}_{t-i} + \boldsymbol{\nu}_t, \quad (12)$$

where $\boldsymbol{\zeta}_1 = \boldsymbol{\Gamma}\overline{\boldsymbol{\beta}}$ and $\boldsymbol{\zeta}_2 = \boldsymbol{\Gamma}\overline{\boldsymbol{\beta}}_1$ are adjustment parameters. This way of writing the VECM is perhaps more intuitive compared to (6) since all the cointegration possibilities (9)-(11) are displayed rather straightforwardly. The representation (12) forms the basis for the VECM formulated in terms of the original $\mathsf{I}(1)$ flow variables $\mathbf{x}_t$:

Theorem 1. ***Error correction model representation for the flow variables.*** *Let $\mathbf{x}_t$ be a p-dimensional vector process of flow variables integrated of order one and multicointegrated such that $r \geq p-r-s > 0$ and such that $\mathbf{X}_t = \sum_{j=1}^{t}\mathbf{x}_j$ is an autoregressive process given by (12). Then $\Delta\mathbf{x}_t$ can be given the following representation:*

$$\begin{aligned}\Delta\mathbf{x}_t &= \boldsymbol{\xi}_1[\boldsymbol{\delta}'\mathbf{Q}_{t-1} - \boldsymbol{\delta}'\boldsymbol{\delta}\boldsymbol{\beta}_2'\mathbf{x}_{t-1}] + \boldsymbol{\xi}_2[\boldsymbol{\delta}_\perp'\mathbf{Q}_{t-1}] \\ &\quad -\boldsymbol{\zeta}_1\Delta\mathbf{Q}_{t-1} - \boldsymbol{\zeta}_2\boldsymbol{\beta}_1{}'\mathbf{x}_{t-1} + \boldsymbol{\Psi}(L)\Delta\mathbf{x}_t + \boldsymbol{\nu}_t\end{aligned} \quad (13)$$

where $\boldsymbol{\xi}_1 = \boldsymbol{\alpha}\overline{\boldsymbol{\delta}}$ and $\boldsymbol{\xi}_2 = \boldsymbol{\alpha}\overline{\boldsymbol{\delta}}_\perp$ are adjustment parameters of dimensions $p\times(p-r-s)$ and $p\times(r-(p-r-s))$, respectively, and $\mathbf{Q}_t = \sum_{j=1}^{t}\boldsymbol{\beta}'\mathbf{x}_j$ is the stock of cumulated I(0) errors. Furthermore $\boldsymbol{\Psi}(L) = \sum_{i=1}^{k-2}\boldsymbol{\Psi}_i L^i$.

Proof. Consider the ECM representation (12) and in place of $\boldsymbol{\alpha}$ use the identity $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\mathbf{P}_\delta + \mathbf{P}_{\delta_\perp})$ where $\mathbf{P}_\delta = \boldsymbol{\delta}(\boldsymbol{\delta}'\boldsymbol{\delta})^{-1}\boldsymbol{\delta}'$ denotes the orthogonal projection matrix on the space spanned by $\boldsymbol{\delta}$. $\mathbf{P}_{\delta_\perp}$ is defined in a similar fashion. The representation now follows straightforwardly by collecting terms and noting that $\mathbf{x}_t = \Delta\mathbf{X}_t$ and $\mathbf{X}_t = \sum_{j=1}^{t}\mathbf{x}_j$ and by using the definition of the stock variables $\mathbf{Q}_t$. ■

In (13) $\boldsymbol{\xi}_1$ adjusts for the $p-r-s$ multicointegrating relations which are given by $[\boldsymbol{\delta}'\mathbf{Q}_{t-1} - \boldsymbol{\delta}'\boldsymbol{\delta}\boldsymbol{\beta}_2'\mathbf{x}_{t-1}]$ and correspond to the 'integral' control mechanisms. $\boldsymbol{\xi}_2$ adjusts for the remaining $r-(p-r-s)$ cointegrating relations amongst the stock variables but these relations do not have the property that they multicointegrate. $\boldsymbol{\zeta}_1$ and $\boldsymbol{\zeta}_2$ adjust for the 'proportional' error correction terms given by $\Delta\mathbf{Q}_t$ and $\boldsymbol{\beta}_1{}'\mathbf{x}_t$. Finally the 'derivative' control mechanisms are given by $\boldsymbol{\Psi}(L)\Delta\mathbf{x}_t$. Note that, ignoring the error term, the flow ECM given in (3) can be obtained as a special case of (13).

The second ECM representation for multicointegrated time series can be stated as follows:

Theorem 2. ***Error correction model representation for the stock variables.*** *Let $\mathbf{x}_t$ be a p-dimensional vector process of flow variables integrated of order one and multicointegrated such that $r \geq p-r-s > 0$ and such that $\mathbf{X}_t = \sum_{j=1}^{t}\mathbf{x}_j$ is an autoregressive process given by (12). Then the process $\Delta\widetilde{\mathbf{x}}_t = (\Delta\mathbf{Q}_t', \boldsymbol{\beta}_1'\mathbf{x}_t, \Delta\mathbf{x}_t'\boldsymbol{\beta}_2)'$ where $\mathbf{Q}_t = \sum_{j=1}^{t}\boldsymbol{\beta}'\mathbf{x}_j$ can be given the following representation:*

$$\begin{aligned}\Delta\widetilde{\mathbf{x}}_t \;=\;& \widetilde{\boldsymbol{\xi}}_1[\boldsymbol{\delta}'\mathbf{Q}_{t-1} - \boldsymbol{\delta}'\boldsymbol{\delta}\boldsymbol{\beta}_2'\mathbf{x}_{t-1}] + \widetilde{\boldsymbol{\xi}}_2[\boldsymbol{\delta}_\perp'\mathbf{Q}_{t-1}] \\ &+\widetilde{\boldsymbol{\zeta}}_1\Delta\mathbf{Q}_{t-1} + \widetilde{\boldsymbol{\zeta}}_2\boldsymbol{\beta}_1'\mathbf{x}_{t-1} + \mathbf{M}\boldsymbol{\Psi}(L)\overline{\boldsymbol{\beta}}_2\Delta\boldsymbol{\beta}_2'\mathbf{x}_{t-1} \\ &+\widetilde{\boldsymbol{\Psi}}(L)\Delta^2\widetilde{\mathbf{x}}_t + \widetilde{\boldsymbol{\nu}}_t\end{aligned} \tag{14}$$

where the parameters and symbols are defined in the following way: $\mathbf{M} = (\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)'$, $\widetilde{\boldsymbol{\xi}}_1 = \mathbf{M}\boldsymbol{\xi}_1$, $\widetilde{\boldsymbol{\xi}}_2 = \mathbf{M}\boldsymbol{\xi}_2$, $\widetilde{\boldsymbol{\zeta}}_1 = (\boldsymbol{\iota}_r - \mathbf{M}\boldsymbol{\zeta}_1)$, $\widetilde{\boldsymbol{\zeta}}_2 = (\boldsymbol{\iota}_s - \mathbf{M}\boldsymbol{\zeta}_2)$, $\boldsymbol{\iota}_r = (\mathbf{I}_r, \mathbf{0}, \mathbf{0})'$, $\boldsymbol{\iota}_s = (\mathbf{0}, \mathbf{I}_s, \mathbf{0})'$, $\widetilde{\boldsymbol{\Psi}}(L) = \mathbf{M}\boldsymbol{\Psi}(L)\mathbf{M}^{-1}\mathbf{D}_\perp(1)$, and $\widetilde{\boldsymbol{\nu}}_t = \mathbf{M}\boldsymbol{\nu}_t$. The dimensions of the '$\sim$' parameters are the same as the parameters without the '$\sim$'.

Proof. We adopt the same technique as in Mellander *et al.* (1992) and define the matrices

$$\mathbf{M} = \begin{pmatrix}\boldsymbol{\beta}' \\ \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_2'\end{pmatrix}, \quad \mathbf{D}(L) = \begin{pmatrix}\boldsymbol{\Delta}_r & & \\ & \boldsymbol{\Delta}_s & \\ & & \mathbf{I}_{p-r-s}\end{pmatrix}, \quad \mathbf{D}_\perp(L) = \begin{pmatrix}\mathbf{I}_r & & \\ & \mathbf{I}_s & \\ & & \boldsymbol{\Delta}_{p-r-s}\end{pmatrix}.$$

Observe that $\mathbf{M}^{-1} = (\overline{\boldsymbol{\beta}}', \overline{\boldsymbol{\beta}}_1', \overline{\boldsymbol{\beta}}_2')$ and $\mathbf{D}(L)\mathbf{D}_\perp(L) = \boldsymbol{\Delta}_p$.

We want to transform the variables in (12) such that error correction adjustments are defined in terms of $\boldsymbol{\beta}'\mathbf{X}_t, \boldsymbol{\beta}_1'\mathbf{X}_t$, and $\boldsymbol{\beta}_2'\Delta\mathbf{X}_t$, which are all I(1), c.f. (9)-(11). Eqn.

(12) can be written as follows:

$$\begin{aligned}\Delta\mathbf{D}(L)\mathbf{D}_{\perp}(L)\mathbf{M}\mathbf{X}_t &= \mathbf{M}\boldsymbol{\alpha}\left[\mathbf{I},\mathbf{0},-\boldsymbol{\delta}\right]\mathbf{D}_{\perp}(L)\mathbf{M}\mathbf{X}_{t-1} \\ &\quad -\mathbf{M}(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,\mathbf{0})\Delta\mathbf{D}_{\perp}(L)\mathbf{M}\mathbf{X}_{t-1} \\ &\quad +\mathbf{M}\boldsymbol{\Psi}(L)\Delta\mathbf{M}^{-1}\mathbf{D}(L)\mathbf{D}_{\perp}(L)\mathbf{M}\mathbf{X}_{t-1}+\mathbf{M}\boldsymbol{\nu}_t.\end{aligned} \tag{15}$$

Now, defining $\widetilde{\mathbf{x}}_t = (\mathbf{X}_t'\boldsymbol{\beta},\mathbf{X}_t'\boldsymbol{\beta}_1,\Delta\mathbf{X}_t'\boldsymbol{\beta}_2)'$ and rewriting (15) yields

$$\begin{aligned}\Delta\widetilde{\mathbf{x}}_t &= \mathbf{M}\boldsymbol{\alpha}\left[\mathbf{I},\mathbf{0},-\boldsymbol{\delta}\right]\widetilde{\mathbf{x}}_{t-1} \\ &\quad +\left\{\mathbf{D}_{\perp}(1)-\mathbf{M}(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,\mathbf{0})+\mathbf{M}\boldsymbol{\Psi}(L)\mathbf{M}^{-1}\mathbf{D}(1)\right\}\Delta\widetilde{\mathbf{x}}_{t-1} \\ &\quad +\mathbf{M}\boldsymbol{\Psi}(L)\mathbf{M}^{-1}\mathbf{D}_{\perp}(1)\Delta^2\widetilde{\mathbf{x}}_{t-1}+\mathbf{M}\boldsymbol{\nu}_t.\end{aligned} \tag{16}$$

Alternatively, by letting $\boldsymbol{\iota}_r = (\mathbf{I}_r,\mathbf{0},\mathbf{0})'$ and $\boldsymbol{\iota}_s = (\mathbf{0},\mathbf{I}_s,\mathbf{0})'$ and collecting terms we obtain

$$\begin{aligned}\Delta\widetilde{\mathbf{x}}_t &= \mathbf{M}\boldsymbol{\alpha}\left[\mathbf{I},\mathbf{0},-\boldsymbol{\delta}\right]\widetilde{\mathbf{x}}_{t-1} \\ &\quad +\left\{\left(\boldsymbol{\iota}_r-\mathbf{M}\boldsymbol{\zeta}_1\right),\left(\boldsymbol{\iota}_s-\mathbf{M}\boldsymbol{\zeta}_2\right),\left(\mathbf{M}\boldsymbol{\Psi}(L)\overline{\boldsymbol{\beta}}_2\right)\right\}\Delta\widetilde{\mathbf{x}}_{t-1} \\ &\quad +\mathbf{M}\boldsymbol{\Psi}(L)\mathbf{M}^{-1}\mathbf{D}_{\perp}(1)\Delta^2\widetilde{\mathbf{x}}_{t-1}+\mathbf{M}\boldsymbol{\nu}_t.\end{aligned} \tag{17}$$

With the definitions of symbols given in the Theorem and by using the identity $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\mathbf{P}_{\delta}+\mathbf{P}_{\delta_{\perp}})$, like in the proof of Theorem 1, the representation (14) results. ∎

The interesting thing to note about Theorem 2 compared to Theorem 1 is that although the same equilibrium relations appear in the ECM, the model is formulated in terms of stock and flow variables rather than in terms of the original flow variables $\mathbf{x}_t$. But of course the adjustment coefficients will be different. The stock variables in the representation (14) are given by the variables $\mathbf{Q}_t = \sum_{j=1}^{t}\boldsymbol{\beta}'\mathbf{x}_j$ and $\mathbf{Q}_t^* = \sum_{j=1}^{t}\boldsymbol{\beta}_1'\mathbf{x}_j$ whereas the flow variables enter through the differenced I(2) trends, $\boldsymbol{\beta}_2'\mathbf{x}_t$. Note, however, that the stock variables $\mathbf{Q}_t^*$ only enter through their derivatives, the flow counterparts $\Delta\mathbf{Q}_t^*$. The ECM given in (2) can be obtained as a special case of (14).

Of course the above representations can be extended to the case where deterministics are included in the models. However, since this complicates the analysis considerably without gaining much additional insight it is excluded in the present exposition. The way that deterministics affect the representation of the I(2) VAR and the statistical analysis is discussed in Paruolo (1996) and Jørgensen *et al.* (1996).

5. An application to US housing construction data

In this section we use I(2) techniques to analyze multicointegration in the US housing data set previously examined by Lee (1992, 1996). The series are from Citibase and are: new privately owned housing units started, $x_t = HSFR$, and new privately owned housing units completed, $y_t = HCP$. Data are monthly seasonally adjusted and covers the period 1968:1-1994:12. In figure 1 x_t and y_t are displayed. Unit root tests (not reported) strongly indicate the presence of a single unit root in both series. Since the $\mathsf{I}(2)$ analysis is based on the cumulated series $X_t = \sum_{j=1}^{t} x_j$ and $Y_t = \sum_{j=1}^{t} y_j$ the graphs of these time series are displayed as well in figure 1. Note that assuming x_t and y_t to be $\mathsf{I}(1)$, X_t and Y_t will be $\mathsf{I}(2)$ by their construction and will measure the stocks of started and completed housing units, respectively.

This is a bivariate example and naturally this simplifies the analysis somewhat. If x_t and y_t are multicointegrated the theory predicts that $s = 0$ (no $\mathsf{I}(1)$ common trends), $r = 1$, (one multicointegrating relation), and $p - r - s = 1$ (a single common $\mathsf{I}(2)$ trend).

Figure 1 about here

5.1. Testing for multicointegration. Lee (1992, 1996) defined the stock of housing units under construction as $Q_t = \sum_{j=1}^{t}(.98x_j - y_j) = .98X_t - Y_t$ and he found strong support for Q_t being $\mathsf{I}(1)$ and cointegrating with y_t, that is x_t, y_t are multicointegrated. However, since Q_t is a constructed variable *a priori* his analysis did not explicitly involve $\mathsf{I}(2)$ variables, *c.f.* the discussion in section 2. Here we adopt the $\mathsf{I}(2)$ analysis both by single equation tests and by tests based on the VAR.

Single equation tests. The single equation procedure has recently been suggested by Engsted *et al.* (1997) as an extension of the Engle and Granger (1987) two-step procedure to the case with $\mathsf{I}(2)$ variables, see also Haldrup (1994). The idea is in the first step to conduct the regression

$$Y_t = \widehat{\kappa}_0 X_t + \widehat{\kappa}_1 \Delta Y_t + \widehat{\delta}_0 + \widehat{\delta}_1 t + \widehat{u}_t \tag{18}$$

and in the second step to test the integration order of $\widehat{u}_t$. For instance, an augmented Dickey Fuller test (or Phillips (1987) $Z-$test) can be conducted. Given that y_t and x_t cointegrate to $\mathsf{I}(0)$ level the critical values for the ADF test for the null of no multicointegration are reported in Engsted *et al.* (1997) for the case where an intercept plus a trend are included in the first step regression as in (18). For $m_1 = 1$ $\mathsf{I}(1)$ regressor and $m_2 = 1$ $\mathsf{I}(2)$ regressor the asymptotic 1% critical value is given by -4.73. Note that the reason for including a trend regressor in (18) is to account for the fact that the generated series X_t and Y_t are likely to drift when x_t and y_t have a non-zero mean. The ADF test gives a test value of -4.82 when 3 lags are included in the auxiliary DF regression[6]. Hence there is rather strong evidence in favour of multicointegration.

[6] The single equation analyses were conducted by using PcGive 9.0, see Hendry and Doornik (1996).

The estimated cointegrating relation (apart from deterministics) is given by

$$Y_t = .947X_t - 8.153\Delta Y_t, \tag{19}$$

so the estimate of κ_0, the percentage of housing constuction starts that are eventually completed, is given by 94.7%.

Alternatively, since cointegration has already been found it is also possible to estimate κ_0 from a more complex dynamic regression including lags of ΔY_t and ΔX_t. The implied 'static' long run solution (see Hendry and Doornik, 1996, p. 234), from a model including 4 lags of ΔY_t and ΔX_t reads:

$$Y_t = .947X_t - 5.581\Delta Y_t - 2.651\Delta X_t \tag{20}$$

and hence it is seen that the estimate of κ_0 happens to be identical to the estimate based on the regression (19).

Tests based on the I(2) VAR. Single equation analysis is not the most efficient way of estimating κ_0. The I(2) VAR model for the variables Y_t and X_t is an alternative and more comprehensive benchmark for analyzing the multicointegrating properties of the data. The statistical analysis for the I(2) model in VAR systems is discussed in Johansen (1995, 1997). The basic model we consider is of the form (6) but we also want to allow for an intercept and a deterministic trend in the model. However, since we do not want the model to generate quadratic trends we restrict the trend to lie in the cointegration space. Essentially this corresponds to a multicointegrating relation with a trend similar to (18). The statistical analysis for this situation is discussed in Jørgensen *et al.* (1996). In the subsequent analysis this restriction has been imposed in the estimation[7].

In order to determine the integration indeces an I(2) VAR model with 12 lags and a trend restricted to lie in the cointegration space was estimated for the variables Y_t and X_t. In table 1 various univariate and multivariate diagnostics of the VAR are reported. As seen there are generally no problems about autocorrelation but due to heteroscedasticity and excess kurtosis there are problems with the normality tests. A possible remedy of this problem is to log transform the variables. However, this is not strictly permitted when the focus is on the stock-flow relations. Since Gonzalo (1994) has found that the Johansen ML procedure is rather robust to discrepancies from the model assumptions due to heteroscedasticity we ignore this problem. In figure 2 the roots of the companion matrix of the VAR(12) are displayed for the case where no I(2) trends are allowed for in the estimation. As seen there is evidence of at

[7]All I(2) VAR computations have been calculated from the the I(2) extension of the CATS in RATS programme. The programme is written by Clara M. Jørgensen and is downloadable from http://www.estima.com/procs/i2index.htm.

Table 1. Misspecification tests of the VAR(12).

Multivariate tests:			
Residual autocorrelation	LM(1)		$\chi^2(4) = 2.79$
Residual autocorrelation	LM(4)		$\chi^2(4) = 7.31$
Normality	LM		$\chi^2(4) = 19.31^{***}$
Univariate tests:		ΔY_t	ΔX_t
ARCH(12)		5.53^{***}	33.33^{***}
Jarque-Bera, Normality		12.56^{***}	7.19^{***}
Skewness		.47	-.13
Excess kurtosis		3.91	3.69
R^2		.95	.91

Note: (*), (**), and (***) denote significance at 1,5, and 10 % levels, respectively.

least two unit roots in the data. Formal testing for the number of unit roots and the integration indeces can be based on the $S_{r,s}$ trace statistics discussed by Johansen (1995, 1997). The results are reported in table 2. This table should be read from the upper left corner (the most restricted model, $r = 0$, $p - r - s = 2$) to the right, and down, until the first hypothesis that cannot be rejected is identified. That is, the hypotheses should be tested successively less and less restricted.

Figure 2 about here

As seen the first hypothesis that cannot be rejected corresponds to the hypothesis where $r = 1$ and $p - r - s = 1$. Hence the tests indicate that there is one common I(2) trend in the data and one polynomially cointegrating (or multicointegrating) relation. Note that $s = 0$, so there are no I(1) trends in the model.

The error correction models. The ECM's presented in theorems 1 and 2 become somewhat simplified in the present situation when there are no I(1) trends and $r = p - r - s$. In this case the I(0) relation is also the multicointegrating relation. The estimates of the model parameters are tabulated in table 3. Note that since $\widehat{\boldsymbol{\delta}}_{\perp} = 0$ and $\widehat{\beta}_1 = 0$ many of the error correction terms will vanish.

As seen the estimated stock of housing units under construction is given by $Q_t = -\boldsymbol{\beta}'(Y_t, X_t) = \sum_{j=1}^{t}(.972x_j - y_j)$. This is only slightly different from Lee's (1992) estimate that 2% of housing starts were never completed. In figure 3 two different measures of the stock variable are displayed and it is apparent that the estimated

Table 2. Joint tests of the integration indeces.
US Housing data set, 1968:1-1994:12.

$p-r$	r	$S_{r,s}$		Q_r
2	0	76.90***	32.72*	22.55*
		47.60	*34.36*	*25.43*
1	1		**13.90**	4.20
			19.87	*12.49*
$p-r-s$		2	1	0

Note: (*), (**), and (***) denote significance at 1,5, and 10 % levels, respectively. Numbers in italics are 95 per cent quantiles, (Jørgensen et al. (1996), table 4). r and $p-r-s$ are the number of $\mathsf{I}(0)$ and $\mathsf{I}(2)$ components.

Table 3. Parameter estimates of the $\mathsf{I}(2)$ VAR for $r=1, s=0$, and $p-r-s=1$.

Multicointegrating relation	$\widehat{\boldsymbol\beta}' = (1, -.972)$, $\widehat{\boldsymbol\delta} = 3.908$, $(\widehat{\boldsymbol\delta}_\perp = 0)$
$\mathsf{I}(2)$ trend	$\widehat{\boldsymbol\beta}_2' = (1, 1.029)$
$\mathbf{M}$−matrix	$\widehat{\mathbf{M}} = (\widehat{\boldsymbol\beta}, \widehat{\boldsymbol\beta}_2)'$
Multicointegrating parameters	$\widehat{\boldsymbol\delta}\widehat{\boldsymbol\beta}_2' = (3.908, 4.022)$
Adjustment coefficients	$\widehat{\boldsymbol\alpha}' = (-.026, .034)$
Gamma matrix	$\widehat{\boldsymbol\Gamma} = \begin{pmatrix} .372 & -.158 \\ -.207 & -.064 \end{pmatrix}$
Derived estimators	
Flow ECM	Stock ECM
$\widehat{\boldsymbol\xi}_1\widehat{\boldsymbol\delta} = \widehat{\boldsymbol\alpha}\widehat{\widehat{\boldsymbol\delta}}\widehat{\boldsymbol\delta} = \widehat{\boldsymbol\alpha}$	$\widehat{\widehat{\boldsymbol\xi}}_1\widehat{\boldsymbol\delta} = \widehat{\mathbf{M}}\widehat{\boldsymbol\xi}_1\widehat{\boldsymbol\delta} = \widehat{\mathbf{M}}\widehat{\boldsymbol\alpha} = \begin{pmatrix} -.059 \\ .009 \end{pmatrix}$
$\widehat{\boldsymbol\xi}_2 = \widehat{\boldsymbol\alpha}\widehat{\widehat{\boldsymbol\delta}}_\perp = 0$	$\widehat{\widehat{\boldsymbol\xi}}_2 = \widehat{\mathbf{M}}\widehat{\boldsymbol\xi}_2 = 0$
$\widehat{\boldsymbol\zeta}_1 = \widehat{\boldsymbol\Gamma}\widehat{\widehat{\boldsymbol\beta}} = \begin{pmatrix} .270 \\ -.074 \end{pmatrix}$	$\widehat{\widehat{\boldsymbol\zeta}}_1 = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \widehat{\mathbf{M}}\widehat{\boldsymbol\zeta}_1\right) = \begin{pmatrix} .657 \\ -.194 \end{pmatrix}$
$\widehat{\boldsymbol\zeta}_2 = \widehat{\boldsymbol\Gamma}\widehat{\boldsymbol\beta}_1 = 0$	$\widehat{\widehat{\boldsymbol\zeta}}_2 = -\widehat{\mathbf{M}}\widehat{\boldsymbol\zeta}_2 = 0$

stock variable rather than the difference $Q_t = X_t - Y_t$ seems to behave more like the original flow variables x_t and y_t. The variable $Q_t = X_t - Y_t$ tends to drift away compared to the flow series. This drift captures the cumulation of starts that are never completed. It is likely that the drift can be partially corrected for by allowing a trend in the cointegration space as we have done in the present study, however, this will only capture part of the discrepancy since the estimate of κ_0 is still strictly less than one[8].

Figure 3 about here

The multicointegrating relation can be written as follows (apart from the deterministic components), compare (9):

$$\widehat{\boldsymbol{\beta}}'(Y_t, X_t)' - \widehat{\boldsymbol{\delta}}\widehat{\boldsymbol{\beta}}_2'(\Delta Y_t, \Delta X_t)' = Y_t - .972X_t + 3.908\Delta Y_t + 4.022\Delta X_t \qquad (21)$$

and the $\mathsf{I}(2)$ trend is given by

$$\boldsymbol{\beta}_2'(Y_t, X_t)' = Y_t + 1.029X_t. \qquad (22)$$

Note that the $\mathsf{I}(2)$ trend is close to being the sum (average) of the two series since κ_0 is not too far away from unity.

The flow ECM can be written as follows:

$$\begin{aligned} \Delta \begin{pmatrix} y_t \\ x_t \end{pmatrix} &= \begin{pmatrix} -.026 \\ .034 \end{pmatrix} (Q_{t-1} - 3.908y_{t-1} - 4.022x_{t-1}) \\ &\quad + \begin{pmatrix} -.270 \\ .074 \end{pmatrix} (\Delta Q_{t-1}) + \text{lags of } \{\Delta y_t, \Delta x_t, \} + \text{constant} + \text{trend}. \end{aligned} \qquad (23)$$

With respect to the stock adjustment ECM this can be written:

$$\begin{aligned} \Delta \begin{pmatrix} Q_t \\ \boldsymbol{\beta}_2'(Y_t, X_t)' \end{pmatrix} &= \begin{pmatrix} -.059 \\ .009 \end{pmatrix} (Q_{t-1} - 3.908y_{t-1} - 4.022x_{t-1}) \\ &\quad + \begin{pmatrix} .657 \\ -.194 \end{pmatrix} (\Delta Q_{t-1}) + \text{lags of } \{\Delta y_t, \Delta x_t, \} + \text{constant} + \text{trend}. \end{aligned} \qquad (24)$$

Of course, from an economic perspective the first row in the stock ECM is most interesting. For both ECM's it is seen that the adjustment coefficients have the correct signs but adjustment also seems to be extremely slow which reflects high costs of adjustment.

[8]Notwithstanding, it should be mentioned that a LR test for the hypothesis that $\beta' = (1, -1)$ cannot be rejected. The test value is 1.36 which should be compared with the $\chi^2(1)$ distribution; the p-value is .24. It is likely that the trend in the cointegration space gives rise to this result.

6. Concluding comments

In this paper we have demonstrated how two separate, though equivalent, error correction model representations can be given to time series that are multicointegrated in the sense of Granger and Lee (1989, 1990). The I(1) model is not adequate for describing multicointegrated time series, so much of the analysis has focused on the I(2) model and how this can be reparametrized to describe the particular features of multicointegrated time series. Through an empirical application of US housing construction data we also demonstrated how the I(2) analysis can be useful in practice as a tool for analyzing multicointegrated time series even though the basic economic variables are really I(1). To the best of our knowledge this provides the first empirical example of explicitly adopting I(2) procedures to multicointegrated time series in the Granger-Lee meaning of the concept. What remains is to consider multicointegration in more advanced models where a potentially larger information set is used to describe the stock building and flow decisions of the housing market. At least the I(2) procedures seem to work well for small stylized models.

References

[1] Campbell, J. Y., (1987). Does saving anticipate declining labour income ? An alternative test of the permanent income hypothesis. *Econometrica*, **55**, 1249-1273.

[2] Engle, R. F., and C.W.J. Granger, (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica* **55**, 2, 251-276.

[3] Engle, R. F. and B. S. Yoo, (1991). Cointegrated economic time series: A survey with new results. In C.W.J. Granger, and R. F. Engle (Eds.). *Long-run economic relations. Readings in cointegration.* Oxford: Oxford University Press.

[4] Engsted, T., and S. Johansen, (1997). Granger's representation theorem and multicointegration. EUI working paper, ECO No. 97/15, European University Institute, Florence.

[5] Engsted, T., J. Gonzalo, and N. Haldrup, (1997). Testing for multicointegration. *Economics Letters*, forthcoming.

[6] Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics,* **60,** 203-233.

[7] Granger, C. W. J., (1983). Cointegrated variables and error correction models. Discussion paper, University of California, San Diego.

[8] Granger, C. W. J., and T. Lee, (1989). Investigation of production, sales and inventory relations using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics*, **4** (supplement), S145-S159.

[9] Granger, C. W. J., and T. Lee, (1990). Multicointegration. In *Advances in Econometrics.*, Vol. **8**, ed. by G. F. Rhodes and T. B. Fomby, Greenwich, CT: JAI Press, 71- 84.

[10] Haldrup, N., (1994). The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables. *Journal of Econometrics,* **63**, 153-81.

[11] Hendry, D. F. and J. A. Doornik, (1996). *Empirical Econometric Modelling: Using PcGive 9.0 for Windows.* International Thomson Business Press

[12] Hendry, D. F. and T. von-Ungern Sternberg, (1981). Liquidity and inflation effects of consumer expenditure. In A. S. Deaton (ed.), *Essays in the theory and measurement of consumers' behaviour.* Cambridge University Press.

[13] Holt, C.C., F. Modigliani, J. F. Muth, and H. Simon, (1960). *Planning production inventories, and work force.* Prentice Hall, Englewood Cliffs.

[14] Johansen, S., (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica,* **59**, 1551-1580.

[15] Johansen, S., (1992). A representation of vector autorgressive processes integrated of order 2. *Econometric Theory,* **8**, 188-202.

[16] Johansen, S., (1995). A statistical analysis of cointegration for I(2) variables. *Econometric Theory,* **11,** 25-59**.**

[17] Johansen, S., (1997). A likelihood analysis of the I(2) model. *Scandinavian Journal of Statistics*, forthcoming.

[18] Jørgensen, C., H. C. Kongsted, and A. Rahbek, (1996). Trend-stationarity in the I(2) cointegration model. Discussion paper 96-12, Institute of Economics, University of Copenhagen.

[19] Kwakernaak, H., and R. Sivan, (1972). *Linear optimal control systems.* Wiley, New York.

[20] Lee, T.-H., (1992). Stock-flow relationships in housing construction. *Oxford Bulletin of Economics and Statistics,* **54**, 419-430.

[21] Lee, T–H., (1996). Stock adjustment for multicointegrated series. *Empirical Economics,* **21**, 633-639.

[22] Mellander, E. A. Vredin, and A. Warne, (1992). Stochastic trends and economic fluctuations in a small open economy. *Journal of Applied Econometrics,* **7**, 369-394.

[23] Metzler, L.A., (1941). The nature and stability of inventory cycles. *Review of Economics and Statistics,* **23**, 113-129.

[24] Paruolo, P., (1996). On the determination of integration indices in I(2) systems. *Journal of Econometrics,* **72**, 313-356.

[25] Paruolo, P., and A. Rahbek, (1996). Weak exogeneity in I(2) systems. Working paper.

[26] Phillips, A. W., (1954). Stabilization policy in the closed economy. *The Economic Journal,* **64**, 290-323.

[27] Phillips, P.C.B., (1987). Time series regression with a unit root. *Econometrica,* **55**, 277-302.

7. FIGURES

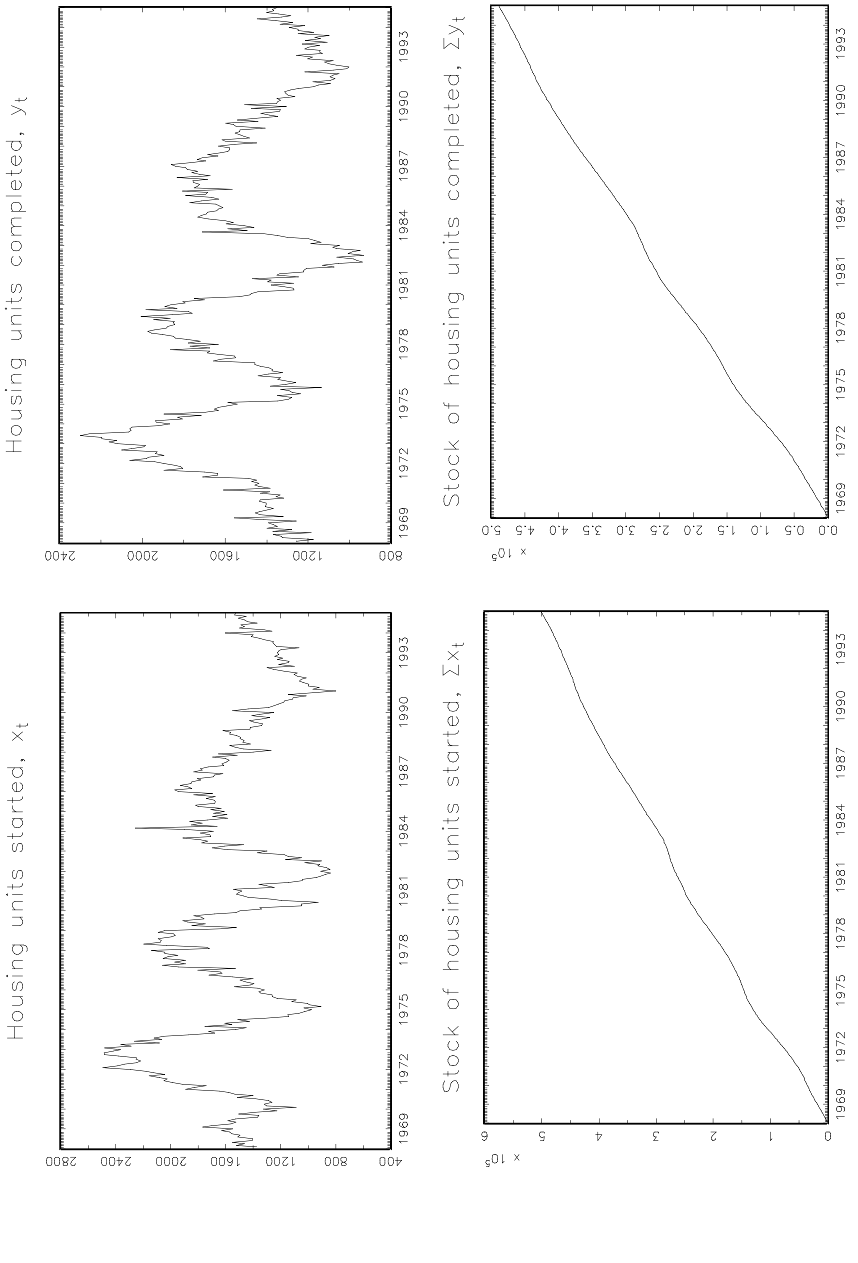

Figure 1: Housing construction data set for US, 1968:1-1994:12, seasonally adjusted. New privately owned housing units started, x_t, and completed, y_t. The corresponding stock series have been generated as $X_t = \sum_{j=1}^{t} x_j$, and $Y_t = \sum_{j=1}^{t} y_j$.

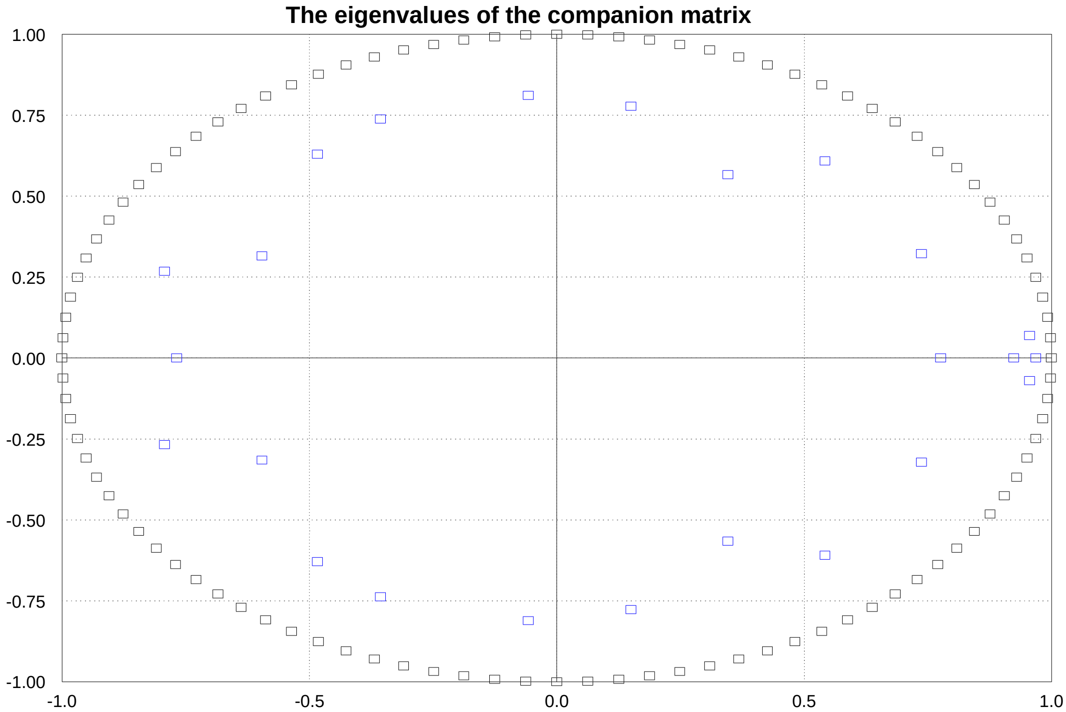

Figure 2: The roots of the companion matrix of the VAR(12).

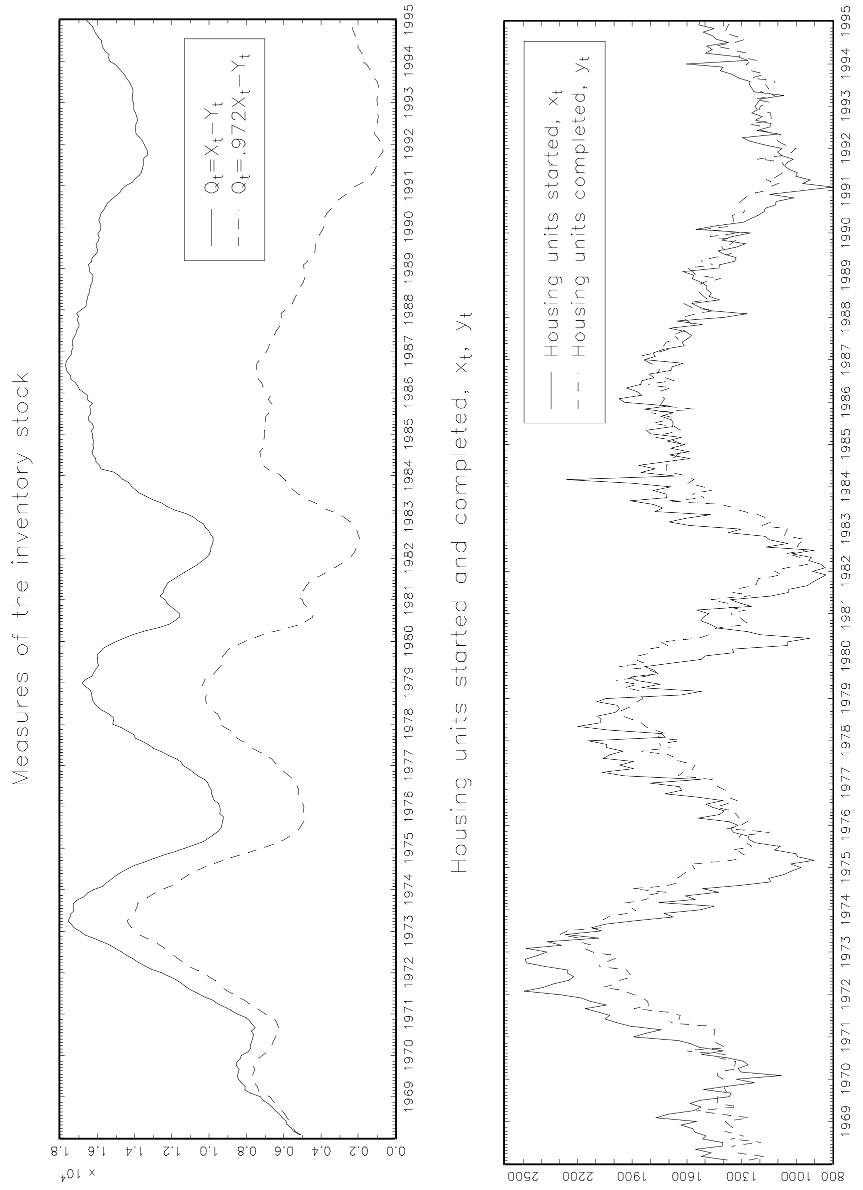

Figure 3: Two different measures of inventories, $Q_t = X_t - Y_t$ and $Q_t = .972X_t - Y_t$ and the associated flow variables x_t and y_t.

Working Paper

1997-5 Torben M. Andersen and Morten Stampe Christensen: Contract Renewal under Uncertainty.

1997-6 Jan Rose Sørensen: Do Trade Unions Actually Worsen Economic Performance?

1997-7 Luca Fanelli: Estimating Multi-Equational LQAC Models with I(1) Variables: a VAR Approach.

1997-8 Bo Sandemann Rasmussen: Long Run Effects of Employment and Payroll Taxes in an Efficiency Wage Model.

1997-9 Bo Sandemann Rasmussen: International Tax Competition, Tax Cooperation and Capital Controls.

1997-10 Toke S. Aidt: Political Internalization of Economic Externalities. The Case of Environmental Policy in a Politico-Economic Model with Lobby Groups.

1997-11 Torben M. Andersen and Bo Sandemann Rasmussen: Effort, Taxation and Unemployment.

1997-12 Niels Haldrup: A Review of the Econometric Analysis of I(2) Variables.

1997-13 Martin Paldam: The Micro Efficiency of Danish Development Aid.

1997-14 Viggo Høst: Better Confidence Intervals for the Population Mean by Using Trimmed Means and the Iterated Bootstrap?

1997-15 Gunnar Thorlund Jepsen and Peter Skott: On the Effects of Drug Policy.

1997-16 Peter Skott: Growth and Stagnation in a Two-Sector Model: Kaldor's Mattioli Lectures.

1997-17 N.E. Savin and Allan H. Würtz: The Effect of Nuisance Parameters on Size and Power; LM Tests in Logit Models.

1997-18 Tom Engsted and Niels Haldrup: Multicointegration in Stock-Flow Models.

CENTRE FOR NON-LINEAR MODELLING IN ECONOMICS

DEPARTMENT OF ECONOMICS - UNIVERSITY OF AARHUS - DK - 8000 AARHUS C - DENMARK
☎ +45 89 42 11 33 - TELEFAX +45 86 13 63 34

Working papers, issued by the Centre for Non-linear Modelling in Economics:

1995-6 David Easley, Nicholas M. Kiefer, Maureen O'Hara and Joseph B. Paperman: Liquidity, Information, and Infrequently Traded Stocks.

1995-13 Tom Engsted, Jesus Gonzalo and Niels Haldrup: Multicointegration and Present Value Relations.

1996-1 Tom Engsted and Niels Haldrup: Estimating the LQAC Model with I(2) Variables.

1996-2 Peter Boswijk, Philip Hans Franses and Niels Haldrup: Multiple Unit Roots in Periodic Autoregression.

1996-3 Clive W.J. Granger and Niels Haldrup: Separation in Cointegrated Systems, Long Memory Components and Common Stochastic Trends.

1996-4 Morten O. Ravn and Martin Sola: A Reconsideration of the Empirical Evidence on the Asymmetric Effects of Money-Supply shocks: Positive vs. Negative or Big vs. Small?

1996-13 Robert F. Engle and Svend Hylleberg: Common Seasonal Features: Global Unemployment.

1996-14 Svend Hylleberg and Adrian R. Pagan: Seasonal Integration and the Evolving Seasonals Model.

1997-1 Tom Engsted, Jesus Gonzalo and Niels Halrup: Testing for Multicointegration.

1997-7 Luca Fanelli: Estimating Multi-Equational LQAC Models with I(1) Variables: a VAR Approach.

1997-12 Niels Haldrup: A Review of the Econometric Analysis of I(2) Variables.

1997-14 Viggo Høst: Better Confidence Intervals for the Population Mean by Using Trimmed Means and the Iterated Bootstrap?

1997-17 N.E. Savin and Allan H. Würtz: The Effect of Nuisance Parameters on Size and Power; LM Tests in Logit Models.

1997-18 Tom Engsted and Niels Haldrup: Multicointegration in Stock-Flow Models.