Central European Journal of Economic Modelling and Econometrics

Cointegration Analysis in the Case of I(2) – General Overview

Michał Majsterek*

Submitted: 1.12.2012, Accepted: 13.03.2013

Abstract

The presented paper aims to analyse both statistical and economic aspects of the model with I(2) variables. The statistical foundations of such models are introduced. The enlargement of possible statistical interpretation is discussed. The economic interpretation of both VECM parameters and common stochastic trends representation is considered in the I(2) domain. The returns of I(2) approach in terms of stock-flows, nominal-real analysis and diasggregation into both long-, short and even medium-run analysis are proved. Potential complications under reflecting I(3) variables are presented.

Keywords: cointegration, I(2) model, VAR

JEL Classification: C32, C51

*University of Łódź, Chair of Econometric Models and Forecasts; e-mail: emfmim@uni.lodz.pl

1 Introduction

Cointegration is most frequently associated with transforming to stationarity the combination of I(1) variables. Basic techniques of estimation both cointegrating vector (two-step Engle and Granger 1987 method) and cointegrating matrix (Johansen 1988 procedure) concerned I(1) categories. Even modified methods of statistical inference were generally connected with integration of order one analysis. Also empirical researches were concentrated on I(1) analysis.
Recently, through Johansen (1994), Johansen (1997) and Paruolo (1996) works, cointegration analysis was enlarged on processes integrated of order two (denoted: I(2)). New, both economic and statistical possibilities of interpretation that models including such variables yield (Juselius 1999) are presented.
The structure of paper is as follows. In section one general features of I(2) processes are considered. There were introduced elements of cointegration analysis inside simple Engle and Granger (1987) approach. Commonly known Dickey and Pantula (1987) test was also discussed with a special emphasis on I(2) analysis. In section two, VAR model for I(2) variables is considered and both statistical and econometric implications result from it. Section three is devoted to economic interpretation of I(2) variables. Expanded with respect to earlier considerations (cf. Majsterek 2008) discussion about relations between I(0), I(1), I(2) shocks and deterministic trends were introduced. In section four potential I(3) model is considered. In section five, comparison between I(0), I(1) and I(2) models was performed. In section six empirical example of I(2) analysis application was presented. Section seven concludes.

2 Features of variables generated by I(2) processes

Process I(2) is by definition second (double) sum of pure random processes (or alternatively: cumulated random walk). Formally, process generating statistic series y is called integrated of order two, if it may be performed as stationary, invertible process ARMA after differencing twice (Engle and Granger 1987).
Discussing the features of variables generated by I(2) it should be remembered that depending on assumed scientific perspective (sample length, data frequency, on which this sample is based) these properties might be changed. These features may be modified depending on the sample length (cf. Figure 1). Irrespective of the chosen horizon of analysis, for I(2) variables impact of past stochastic shock, not only will not terminate (as in the I(0) case), but in the contrast of I(1) processes, will be enforced. It implies a persistence of shock effects for increases and growth rates of variables. Simultaneously (which is important in the context I(2) processes identification) differences between properties of variables generated by I(1) and I(2) processes asymptotically disappear, because sample enlargement make I(2) processes closed to I(1).
In the small sample I(2) processes have features similar to explosive processes, i.e.

nonintegrated (Haldrup 1999). Both in the first and the second case the impact of stochastic shock increases, but for I(2) processes this impact increases slower. The differences between I(2) and explosive processes are crucial. I(2) processes include by definition two unit roots, while explosive processes have no such root. Osiewalski and Pipień (1999) argued, that explosive processes we can treat as stationary with respect to the future shocks, while I(2) processes achieve stationarity only after application difference filter twice. Consequently, explosive processes are often generated by AR(1) processes, in which inertia parameter $|\alpha_1| > 1$, while I(2) processes by autoregressive (or ARMA) process of order at least two.

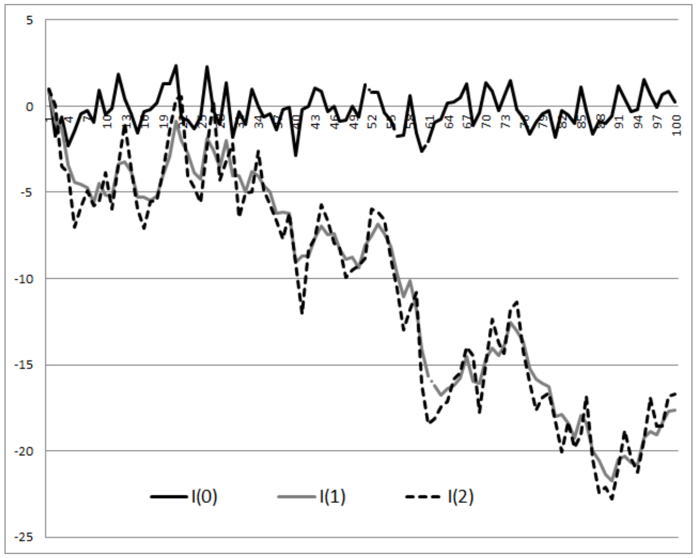
With respect to the variables generated by I(2) processes, it is useful to discriminate strictly long-run shocks and disturbances which are more persistent than transitory, however dominated by long-run I(2) processes. These shocks, often identified with stochastic cycles across stochastic trends I(2), are connected with processes integrated of order one (cf. Figure 1).

Differences interpretation stochastic trends I(1) depending on the highest integration order of variables presented in the model have their source in the domination rule (cf. Figure 1). Process with the highest integration order has always dominant character, hence just this process will has win the stochastic trend role. Process integrated of one order lower $d-1$ will has at most (if $d \geq 2$ stochastic trend I(d) dominates, $d \geq 2$) character of stochastic cyclical, the longer periods the higher is $d-1$ (cf. Juselius 2006). Analysis of Figure 1 allows us to recognise, that if we dispose long sample with low frequency (for example yearly data), then from both estimation and forecasting point of view we do not commit a serious error treating variables as I(1). If research interest is in the cyclical deviations, discrimination between I(1) and I(2) is necessary. In the case of $d = 1$ stochastic cycles are excluded by definition and there are discriminated permanent shocks I(1) and transitory I(0) shocks only.

Very interesting is the integrated processes interpretation related to the basic classification of economic categories. Flows may be treated as first increments of connected with them stock categories, for example inflation is a price increase and chain indices are the increments of fixed base indices. It means, that almost always flows will be integrated of order one level lower than respective stocks. It may be supposed, that nominal categories should be integrated of higher order one level higher than respective real variable. However although the dependency between order of integration for stocks and associated with them flows may be regarded as some regularity (Haldrup 1994), with respect to nominal and real values there are exceptions from this rule. From Haldrup (1999) considerations it follows, that because in economic reality very rarely I(3) processes occur, I(2) variables are rather stocks than flows and more often nominal than real categories, tests results shall be treated with caution if not confirmed these economic suggestions. We may expect, that for flows in actual prices the best interpretable result would be I(1), for nominal stocks: I(2), while for real flows: stationary or I(1).

In the case of the modelling with I(2) variables spurious regressions problem is

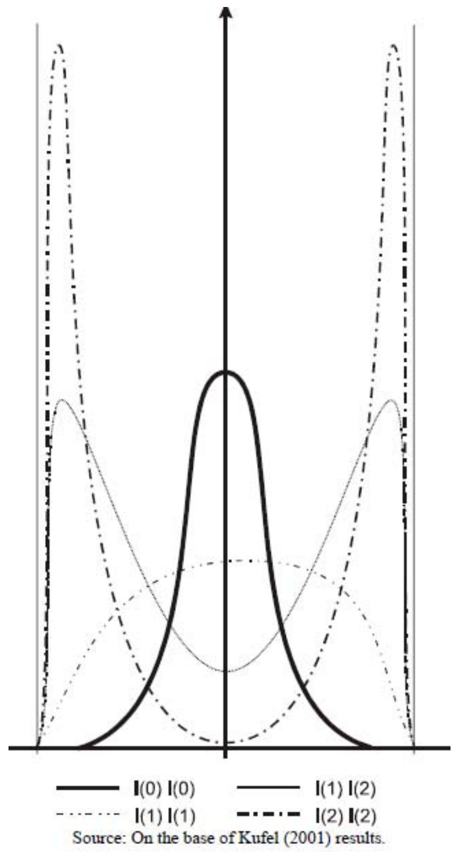
Figure 1: Trends and stochastic cycles



sharpened. Probability of such dependencies identification arises, if all variables are integrated of order two. The distribution of correlation coefficient between two uncorrelated I(2) variables is bimodal (Figure 2), which although still guarantee zero expected value cause, but the variance of bimodal distribution is the reason that not only spurious regressions, misleading, but strictly nonsense dependency may be identified. In this way in I(2) analysis it is necessary to clearly discriminate this two, earlier used interchangeably, terms (Yule 1926, Granger and Newbold 1974). Spurious dependencies were not identified in relationships (noncointegrating by assumption) between on the one hand side stationary variables, and I(1) or I(2) on the other hand. The most interesting Banerjee et. al. (1993) research's results was the confirmation of high probability of the spurious correlation identification between I(1) and I(2) variables. It was rather intuitively expected (from the non-cointegration between I(1) and I(2) processes) the analogy to relationships between I(0) with I(1) or I(2), for which identification spurious regression probability is very small. Kufel's (2001) experiments confirmed, that the true null hypothesis, that I(1) and I(2) variables are uncorrelated, is rejected even more often (in 87.2%) than in the case, when both variables are integrated of the same order I(1) (in this case rejections probability is 78.5%). It means, that with the increase of nonstationarity order, spurious regressions problem arise.

In the case of two I(1) variables cointegration relationship has "timeless", static character. It means, that if x_{1t} and x_{2t} variables are cointegrated, then for example $x_{1,t-1}$ and x_{2t} are cointegrated. It is correct, with respect to I(1) variables to briefly

Figure 2: Test statistic distributions for correlation coefficient in the case variables I(0), I(1) and I(2).



define long-run dependency, apart from time index, i.e. x_1 and x_2. In the case of two variables integrated of order two (cf. Haldrup 1999) and simultaneously cointegrated CI(2,2), i.e. x_{1t} and x_{2t}, variables $x_{1,t-1}$ and x_{2t} are not cointegrated. It results from the fact, that if $x_{1t} - x_{2t} \sim I(0)$, then $x_{1,t-1} - x_{2t} = (x_{1t} - x_{2t}) - \Delta x_{1t} \sim I(1)$, because there no exists cointegration relationship between stationary combination $x_{1t} - x_{2t} \sim I(0)$ and integrated of order one Δx_{1t} variable.
In the case of the variables I(2) modelling, the sense of cointegration is more complicated than for variables I(1). Due to the classical Engle and Granger (1987) definition cointegration means the existence of such nonzero (cointegrating) vector β, which implies that $\beta^T \mathbf{Y}$ is integrated of order $d - b < d$, while $\mathbf{Y} \sim I(d)$. With respect to multi-dimensional case, it was assumed, that all variables in the matrix $\mathbf{Y}$ are I(d). Johansen 1988 modified this assumption, assuming as d means the

highest of integration orders among variables reflected in the model. Flores and Szafarz (1996)summarized this discussion concerning the definition of cointegration and proposed to call multi-dimensional process including variables integrated of order at most d cointegrated, if there exists nontrivial linear combination of its components, which is integrated of order $d - b < d$. In the case when $d = 1$, combination $\beta^T\mathbf{Y}$ (random term) is integrated of order zero, which implies that all cointegration regression must be stationary. The estimator of vector parameters β (in the multi-dimensional case: cointegrating matrix) is then superconsistent. If however $d = 2$, not all cointegrating dependencies are stationary. Also problem of consistency cointegrating vector estimator is more complex.

Cointegration relationship in I(2) domain mostly (but not always) is the dependency between levels of variables. Usually, it is long-run, however also medium-run cointegration is considered. Synthetic comparison of cointegration in the case I(1) and I(2) is presented in Table 1.

In the case of cointegration between I(2) variables cointegrating vector estimator is super-superconsistent:

$$p\left(\lim\left(T^2(\hat{\beta}-\beta)\right)\right) = 0 \tag{1}$$

which means faster (requiring less long sample) achievement of desired asymptotic features than in the case of cointegration between I(1) variables. Super-superconsistency of estimator reaches not only for cointegration CI(2,2), in which random term is stationary, but also for cointegrating dependencies CI(2,1), in which random term is I(1) (wider considerations in the next section). Cointegration CI(2,1) is interpreted as "uncontrolled" long-run stochastic shocks annihilation is then caused by the centripetal (cointegrating) forces between variables. The shocks disturbing this relationship are more persistent than temporary disturbances from cointegrating relationships CI(2,2), but also vanish in the long-run. In this sense CI(2,1) cointegration is also long-run equilibrium relationships.

Due to Diebold and Nerlove (1990)time series decomposition (modified by Romański and Strzała 1995) it may be distinguished stochastic trend, deterministic trend, and cyclical term also broken down into deterministic and stochastic and pure random term. In the case of model with I(2) variables, I(2) trends may be identified with stochastic trend (cf. Juselius 1999), whereas I(1) trends rather with stochastic cycles across long-run I(2) trends. In this sense I(2) analysis allows to distinguish medium-run deviations, which are absent in the model with I(1) variables.

In the case of testing by DF with respect to I(2) variables very often the true I type error probability is higher (sometimes significantly higher) than nominal level of significance (cf. Dickey and Pantula 1987). Meanwhile just arbitrary assumed nominal (hence wrong) I type error probability is applied to determine critical values. In particular, the probability of the null hypothesis I(1) (against alternative I(0)) rejection will higher in the case, when in fact the series is I(d), where $d \geq 2$, than if the series is indeed I(1). Consequently, against intuition, "more nonstationary" I(2) process will be more often misled with stationary process than the random walk

Table 1: Order of cointegration of variables and estimator properties

Integration order	CI(0,0)	CI(1,1)	CI(2,1)	CI(2,2)	Non-cointegration
I(0)	consistency	–	–	–	–
I(1)	consistency	super-consistency	–	–	non-consistency
I(2)	consistency	super-consistency	super-super-consistency	super-super-consistency	non-consistency

with I(0) processes. Hence, if is supposed that order of integration equals $d \geq 2$, it is proper to apply test, which not verify I(1) against I(0), before checking the hypothesis, that series is I(2) versus I(1). Dickey and Pantula (1987) proposed a test fulfilling above conditions. It is assumed in this test, that all variables are generated by AR(p) process, which after isomorphic transformation has a form:

$$\Delta^p y_t = \theta_1 y_{t-1} + \theta_2 \Delta y_{t-1} + \ldots + \theta_p \Delta^{p-1} y_{t-1} + \varepsilon_t \tag{2}$$

The number of nonsignificantly different from zero θ_j $(j = 1, \ldots, p)$ parameters suggests the integration order of variable. Applying significance tests t we may verify a hypothesis that p unit roots exist, using so called "pseudo-t" (t^*) statistics.
The term $\Delta^p y_t$ the most significantly depends on $\Delta^{p-1} y_{t-1}$ (it may be performed, that if $\theta_p = 0$, then $\theta_1 = \ldots = \theta_{p-1} = 0$). If test results not allow us to reject the hypothesis $\theta_p = 0$ against $\theta_p < 0$, then there are no bases to reject $\theta_j = 0$ ($j = 1, \ldots, p-1$). It may be in this case assumed, that series is I(p), because has p unit roots. From the identification of stochastic shocks I(2) point of view, it is purposeful to choose as initial $p = 2$ (in practice variables integrated of order higher than 2 rather not occur). Then (2) may be simplified to isomorphic transformation of AR(2)

$$\begin{aligned}
\Delta^2 y_t &= (\alpha_1 - 1) y_{t-1} - \Delta y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t \\
&= (\alpha_1 - 2) y_{t-1} + (\alpha_2 + 1) y_{t-2} + (\alpha_2 + 1) y_{t-1} - (\alpha_2 + 1) y_{t-1} + \varepsilon_t \\
&= (\alpha_1 + \alpha_2 - 1) y_{t-1} - (\alpha_2 + 1) \Delta y_{t-1} + \varepsilon_t
\end{aligned}$$

Variable is I(2), if in the AR(2) model $\alpha_1 = 2 \wedge \alpha_2 = -1$. Then, because from (2) and (3) results, that:

$$\theta_1 = \alpha_1 + \alpha_2 - 1, \ \theta_2 = -(\alpha_2 + 1) \tag{3}$$

so null hypothesis $\theta_1 = \theta_2 = 0$ is equivalent with $H_0 : y \sim I(2)$. Its rejection means, that variable is integrated of order at most one. In this case the next step shall be initialized to verify $I(p-1)$ against $I(p-2)$ (in our example: I(1) vs. I(0)). The null hypothesis should be rejected, if value of this t^* statistic is significantly smaller than zero. Verification procedure starts from the assumption about potentially the highest

integration order.
Dickey and Pantula (1987) procedure has some disadvantages (wider discussion in: Majsterek 2008). It were connected with applying "from specific to general" strategy, similarly as other unit root tests, test DP has low power. Additionally, the analysis based on Dickey and Pantula statistic has mechanic, non-system character.
The problem of traditional cointegration testing (for example application of cointegration Engle and Granger test) also complicates for variables integrated of order two. In I(2) domain no bases to reject the null hypothesis is not equivalent with the absence of cointegration between variables, but implies lack of the speediest cointegrating CI(2,2) dependencies only. Only no bases to reject null hypothesis, that first increments of disturbances are nonstationary, does imply that there are no cointegrating relationships CI(2,1), and the absence of long-run dependencies between variables. Integration order of residuals equal one, as results from Table 2, may (but not necessarily must) suggests CI(2,1) cointegration, then the more proper approach in cointegration testing in the I(2) processes case is the analysis based on VAR models.

Table 2: Comparison of cointegration relationships between I(1) and I(2) variables.

I(1) variables	I(2) variables
static relationship, "timeless"	"static" and "dynamic" relationship, not always "timeless"
stationary relationship CI(1,1)	stationary CI(2,2) and nonstationary CI(2,1) relationship
relationship between levels of variables	relationship between levels of variables, sometimes additionally between levels and differences
long-run equilibrium relationship	long-run equilibrium CI(2,1) relationship, long- and medium-run equilibrium CI(2,2) relationship, medium-run CI(1,1) relationship

3 VECM model in the case of I(2) variables

Let us consider vector autoregression model (VAR):

$$\mathbf{Y}_t = \mathbf{\Pi}^{(1)}\mathbf{Y}_{t-1} + \mathbf{\Pi}^{(2)}\mathbf{Y}_{t-2} + \ldots + \mathbf{\Pi}^{(S)}\mathbf{Y}_{t-S} + \mathbf{\Sigma}_t \qquad (4)$$

where: $\mathbf{Y}_{t-s}$ – matrix of observation on variables in the model in the period $t-s$, values of these variables for $t < 0$ are assumed as non-random and predetermined $\mathbf{\Pi}^{(s)}$ – $M \times M$ matrix of parameters.
This model may be transformed to VECM (vector error correction model) form:

$$\Delta\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \sum_{s=1}^{S-1}\mathbf{\Gamma}_s\Delta\mathbf{Y}_{t-s} + \mathbf{\Sigma}_t \qquad (5)$$

where $\mathbf{\Pi} = \sum_{s=1}^{S} \mathbf{\Pi}^{(s)} - \mathbf{I}$ and $\mathbf{\Gamma}_s = \sum_{j=s+1}^{S} \mathbf{\Pi}^{(j)}$.
Analogously, after isomorphic transformation:

$$\Delta^2 \mathbf{Y}_t = \mathbf{\Pi Y}_{t-1} + \mathbf{\Gamma} \Delta \mathbf{Y}_{t-1} + \sum_{s=1}^{S-2} \mathbf{\Psi}_s \Delta^2 \mathbf{Y}_{t-s} + \mathbf{\Sigma}_t \tag{6}$$

where:

$$\mathbf{\Gamma} = \sum_{s=1}^{S-1} \mathbf{\Gamma}_s - \mathbf{I} \tag{7a}$$

$$\mathbf{\Psi}_s = - \sum_{j=s+1}^{S-1} \mathbf{\Gamma}_j \tag{7b}$$

For example, for $S = 3$:

$$\begin{aligned} \Delta^2 \mathbf{Y}_t &= \mathbf{\Pi Y}_{t-1} - \mathbf{Y}_{t-1} + \mathbf{Y}_{t-2} + \mathbf{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{Y}_{t-2} + \mathbf{\Sigma}_t \\ &= \mathbf{\Pi Y}_{t-1} + (\mathbf{\Gamma}_1 - I) \Delta \mathbf{Y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{Y}_{t-2} + \mathbf{\Sigma}_t \\ &= \mathbf{\Pi Y}_{t-1} + (\mathbf{\Gamma}_1 - I) \Delta \mathbf{Y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{Y}_{t-2} + \mathbf{\Gamma}_2 \Delta \mathbf{Y}_{t-1} - \mathbf{\Gamma}_2 \Delta \mathbf{Y}_{t-1} + \mathbf{\Sigma}_t \\ &= \mathbf{\Pi Y}_{t-1} + (\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 - I) \Delta \mathbf{Y}_{t-1} - \mathbf{\Gamma}_2 \Delta^2 \mathbf{Y}_{t-1} + \mathbf{\Sigma}_t \end{aligned} \tag{7c}$$

Matrix $\mathbf{\Gamma}$ $(M \times M)$ is called mean lag matrix.
Alternatively it may be considered representation, which will be isomorphic transformation of (6):

$$\Delta^2 \mathbf{Y}_t = \mathbf{\Pi Y}_{t-S} + \tilde{\mathbf{\Gamma}} \Delta \mathbf{Y}_{t-1} + \sum_{s=1}^{S-2} \mathbf{\Psi}_s \Delta^2 \mathbf{Y}_{t-s} + \mathbf{\Sigma}_t \tag{8}$$

where $\tilde{\mathbf{\Gamma}} = \sum_{s=1}^{S-1} \tilde{\mathbf{\Gamma}}_s - \mathbf{I}$.
In the case of joint stationarity (no nonstationary variables in the system), cointegrating dependencies identification is superfluous, however structuralisation restrictions may be imposed. All representations (4)–(8) may in this case solution with respect to the stationary shocks. This solution is called vector moving average (VMA) representation:

$$\mathbf{Y}_t = \left(\mathbf{I} - \mathbf{\Pi}^{(1)} L - \cdots - \mathbf{\Pi}^{(S)} L^S \right)^{-1} \Sigma_t \tag{9}$$

and may be calculated because $\mathbf{I} - \mathbf{\Pi}^{(1)} L - \cdots - \mathbf{\Pi}^{(S)} L^S$ matrix is nonsingular, if shocks are short-run (according to stationary processes properties roots of $\mathbf{I} - \mathbf{\Pi}^{(1)} L - \cdots - \mathbf{\Pi}^{(S)} L^S$ lie outside unit circle, and so all characteristic roots of this matrix are nonzero).
All shocks influencing variables in the VAR model are in this case transitory and not cumulate (rows of matrix $\mathbf{I} - \mathbf{\Pi}^{(1)} L - \cdots - \mathbf{\Pi}^{(S)} L^S$ are connected with variable, which

is influenced by shock, whereas columns exhibit short-run shocks sources).
If rank of $\mathbf{\Pi}$ (which is the same as rank of $\mathbf{I}-\mathbf{\Pi}^{(1)}L-\cdots-\mathbf{\Pi}^{(S)}L^S$) fulfils $R<M$, then from noninvertibility of matrix $\mathbf{I}-\mathbf{\Pi}^{(1)}L-\cdots-\mathbf{\Pi}^{(S)}L^S$, moving average representation should be replaced by the following model (cf.Engle and Granger 1987, derivation in: Johansen 1995a pp. 40):

$$\mathbf{Y}_t = \mathbf{c}| \sum_{i=1}^{t} \Sigma_i + \mathbf{c}|(L)\mathbf{\Sigma}_t \tag{10}$$

where:

$$\mathbf{c}| = \mathbf{B}_\perp \left(\mathbf{A}_\perp^T \left(\sum_{s=1}^{S-1} \mathbf{\Gamma}_s - \mathbf{I} \right) \mathbf{B}_\perp \right)^{-1} \mathbf{A}_\perp^T,$$

$\mathbf{A}_\perp = [\bar{a}_{ij}]$, $\mathbf{B}_\perp = [\bar{b}_{ij}]$ - $M\times(M-R)$ orthogonal compliments of matrix $\mathbf{A}$ and $\mathbf{B}$ respectively ($\mathbf{A}^T\mathbf{A}_\perp = \mathbf{0}$ and $\mathbf{B}^T\mathbf{B}_\perp = \mathbf{0}$), ($r\left[\begin{array}{cc}\mathbf{A} & \mathbf{A}_\perp\end{array}\right] = M$ and $r\left[\begin{array}{cc}\mathbf{B} & \mathbf{B}_\perp\end{array}\right] = M$).
The form (10) is called common stochastic trends I(1) representation.
VECM has solution in the form (10) only if, $\mathbf{c}| = \mathbf{B}_\perp \left(\mathbf{A}_\perp^T \left(\sum_{s=1}^{S-1} \mathbf{\Gamma}_s - \mathbf{I}\right) \mathbf{B}_\perp\right)^{-1} \mathbf{A}_\perp^T$ matrix exists. It is connected with the implicitly assumed in I(1) analysis assumption about the full rank of matrix $(\mathbf{A}_\perp^T \left(\sum_{s=1}^{S-1} \mathbf{\Gamma}_s - \mathbf{I}\right) \mathbf{B}_\perp)$. It should be noted, that above I(1) model condition was defined both for matrix parameters from the primary representation (VECM), and from the solution (dual representation).
In the case of double reduced rank $r(\mathbf{\Pi}) = R < M$ and $r\left(\mathbf{A}_\perp^T \left(\sum_{s=1}^{S-1} \mathbf{\Gamma}_s - \mathbf{I}\right) \mathbf{B}_\perp\right) =$ $= P_1 < M - R$, the solution in the form of common stochastic trends model is as follows:

$$\mathbf{Y}_t = \mathbf{c}|_1 \sum_{i=1}^{t} \mathbf{\Sigma}_i + \mathbf{c}|_2 \sum_{j=1}^{t} \sum_{i=1}^{j} \mathbf{\Sigma}_i + C(L)\mathbf{\Sigma}_t \tag{11}$$

where $\mathbf{c}|_2$ is parameters matrix connected with stochastic I(2) trends $\sum_{j=1}^{t} \sum_{i=1}^{j} \mathbf{\Sigma}_i$.
The $\bar{\bar{c}}|_{mn}$ element of $M\times M$-dimensional I(2) shocks matrix $\mathbf{c}|_2$ measures the impact of permanent (double cumulative) shock from n-th variable on m-th variable. The $\bar{c}|_{mn}$ element of $M\times M$ -dimensional matrix $\mathbf{c}|_1$ is the measure of medium-run shock impact from n-th variable on m-th variable.
Matrix of long-run shocks $\mathbf{c}|_2$ may be decomposed:

$$\mathbf{c}|_2 = \mathbf{B}_{2\perp}\mathbf{A}_{2\perp}^T \mathbf{\Gamma}\mathbf{B} \left(\mathbf{B}^T\mathbf{B}^{-1} \left(\mathbf{A}^T\mathbf{A}^{-1}\mathbf{A}^T\mathbf{\Gamma} - \sum_{s=1}^{S-2} \mathbf{\Psi}_s \right) \mathbf{B}_{2\perp}^T \right)^{-1} \mathbf{A}_{2\perp}^T \tag{12}$$

where $\mathbf{\Gamma} = \sum_{s=1}^{S-1} \mathbf{\Gamma}_s - I$ and $\Psi_s = -\sum_{j=s+1}^{S-1} \mathbf{\Gamma}_j$.
The $\mathbf{A}_{2\perp}^T$ matrix consists of coefficients of the baseline (independent) stochastic I(2)

trends. Matrix $\mathbf{B}_{2\perp}$ is the key component of the weights matrix for common stochastic I(2) trends, defined as $\mathbf{A}_{2\perp}^T \sum_{j=1}^{t} \sum_{i=1}^{j} \Sigma_i$. Independent common stochastic I(2) trends may be performed as:

$$\bar{\bar{a}}_{1n} \sum \sum \mathbf{\Sigma}_{1t} + \cdots + \bar{\bar{a}}_{Mn} \sum \sum \mathbf{\Sigma}_{Mt}, \quad n = 1, \ldots, P_2 \tag{13a}$$

where $\bar{\bar{a}}_{ij}$ is the element of matrix $\mathbf{A}_{2\perp}$.
Independent common stochastic trends I(1) may be denoted as:

$$\bar{a}_{1n} \sum \mathbf{\Sigma}_{1t} + \cdots + \bar{a}_{Mn} \sum \mathbf{\Sigma}_{Mt}, \quad n = 1, \ldots, P_1 \tag{13b}$$

where $\bar{a}_{ij}$ is the element of matrix $\mathbf{A}_{1\perp}$.
Chosen m-th variable is influenced by combination of common stochastic trends I(2):

$$\begin{gathered} \bar{w}_{m1} \left(\bar{\bar{a}}_{11} \sum \sum \varepsilon_{1t} + \cdots + \bar{\bar{a}}_{M1} \sum \sum \varepsilon_{Mt} \right) + \cdots + \\ + \cdots + \bar{w}_{m,P_2} \left(\bar{\bar{a}}_{1,P_2} \sum \sum \varepsilon_{1t} + \cdots + \bar{\bar{a}}_{M,P_2} \sum \sum \varepsilon_{Mt} \right) \end{gathered} \tag{13}$$

where coefficient $\bar{w}_{mr}$ is the element of weights matrix

$$\tilde{\mathbf{B}}_{2\perp} = \mathbf{B}_{2\perp} \mathbf{A}_{2\perp}^T \mathbf{\Gamma} \mathbf{B} \left(\mathbf{B}^T \mathbf{B}^{-1} \left(\mathbf{A}^T \mathbf{A}^{-1} \mathbf{A}^T \mathbf{\Gamma} - \sum_{s=1}^{S-2} \mathbf{\Psi}_s \right) \mathbf{B}_{2\perp}^T \right)^{-1}$$

connected with m-th variable and r -th independent common stochastic I(2) trends ($r = 1, \ldots, P_2$).
Matrix $\mathbf{A}_{1\perp}^T$ consists of coefficients of independent stochastic I(1) trends, which are defined by the formula (13a). It should be stressed, it is impossible to decompose medium-run matrix of shocks $\mathbf{c}|_1$ analogously to the $\mathbf{c}|_2$ case. Then matrix $\mathbf{B}_{1\perp}$ may not be interpreted as the component of weights matrix connected with common stochastic I(1) trends, because such weights matrix were not defined in the representation (11).
The interpretation of $\mathbf{A}_\perp$ and $\mathbf{B}_\perp$ matrices is clear only in I (1) case. In I(2) analysis, the additional decomposition of stochastic trends $\mathbf{A}_\perp$ and their weights $\mathbf{B}_\perp$ into the matrices defining I(1) and I(2) trends is required. It this purpose the following dependencies are useful:

$$\mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T + \mathbf{A}_\perp \left(\mathbf{A}_\perp^T \mathbf{A}_\perp \right)^{-1} \mathbf{A}_\perp^T = \mathbf{I} \tag{14}$$

$$\mathbf{B} \left(\mathbf{B}^T \mathbf{B} \right)^{-1} \mathbf{B}^T + \mathbf{B}_\perp \left(\mathbf{B}_\perp^T \mathbf{B}_\perp \right)^{-1} \mathbf{B}_\perp^T = \mathbf{I} \tag{15}$$

which allow us to obtain (cf. Paruolo 2000):

$$\bar{\mathbf{A}}_\perp \mathbf{\Xi} = \mathbf{A}_\perp \left(\mathbf{A}_\perp^T \mathbf{A}_\perp \right)^{-1} \mathbf{\Xi} = \mathbf{A}_{1\perp} \tag{16}$$

because:

$$\begin{aligned}\mathbf{A}_{1\perp} &= \left(\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T + \mathbf{A}_{\perp}\left(\mathbf{A}_{\perp}^T\mathbf{A}_{\perp}\right)^{-1}\mathbf{A}_{\perp}^T\right)\mathbf{A}_{1\perp} \\ &= \mathbf{A}_{\perp}\left(\mathbf{A}_{\perp}^T\mathbf{A}_{\perp}\right)^{-1}\left(\mathbf{A}_{\perp}^T\mathbf{A}_{1\perp}\right) \\ &= \bar{\mathbf{A}}_{\perp}\mathbf{\Xi}\end{aligned}$$

and

$$\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{A}_{1\perp} = \mathbf{0}$$

Analogously:

$$\bar{\mathbf{B}}_{\perp}\mathbf{N} = \mathbf{B}_{\perp}\left(\mathbf{B}_{\perp}^T\mathbf{B}_{\perp}\right)^{-1}\mathbf{N} = \mathbf{B}_{1\perp}, \tag{17}$$

$(M-R)\times P_1$-dimensional (where $P_1 < M-R$) matrices $\mathbf{\Xi}$, $\mathbf{N}$ fulfil from the reduced rank condition $\left(\mathbf{A}_{\perp}^T\left(\sum_{s=1}^{S-1}\mathbf{\Gamma}_s - \mathbf{I}\right)\mathbf{B}_{\perp}\right)$ equality:

$$\mathbf{A}_{\perp}^T\mathbf{\Gamma}\mathbf{B}_{\perp} = \mathbf{\Xi}\mathbf{N}^T \tag{18}$$

This is analogous to the famous from I(1) analysis decomposition:

$$\mathbf{\Pi} = \mathbf{A}\mathbf{B}^T \tag{19}$$

where:

$\mathbf{A} = [\alpha_1\alpha_2\ldots\alpha_R]_{MxR}$ - weights matrix (adjustments matrix),

$\mathbf{B} = [\beta_1\beta_2\ldots\beta_R]_{MxR}$ - matrix consisting of independent cointegrating vector.

The main difference is that matrices $\mathbf{\Xi}$ and $\mathbf{N}$ are not directly interpretable. Simultaneously, following dependencies are fulfilled:

$$\mathbf{A}_{\perp}^T\mathbf{A}_{1\perp} = \mathbf{\Xi} \tag{20}$$

$$\mathbf{B}_{\perp}^T\mathbf{B}_{1\perp} = \mathbf{N} \tag{21a}$$

Formulas (16)-(17) allows us to project $\mathbf{B}_{\perp}$ into the medium-run stochastic trends I(1) subspace, whereas projection into I(2) trends subspace is possible by decomposition:

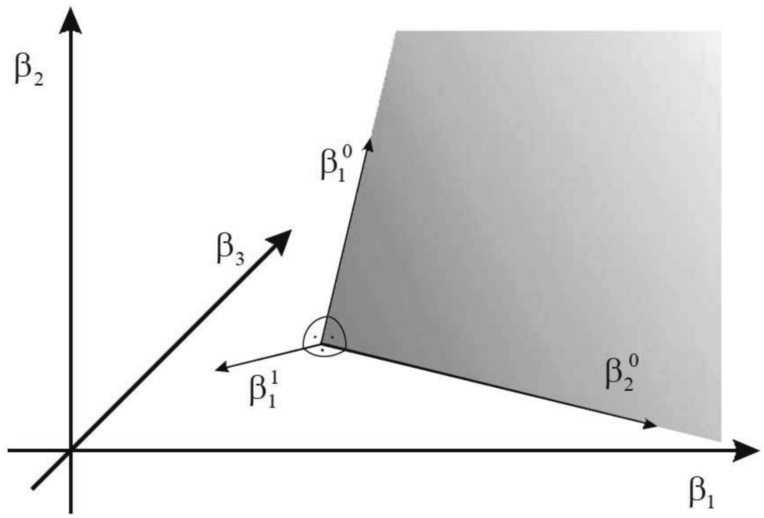
$$\mathbf{A}_{\perp}\mathbf{\Xi}_{\perp} = \mathbf{A}_{2\perp} \tag{21}$$

$$\mathbf{B}_{\perp}\mathbf{N}_{\perp} = \mathbf{B}_{2\perp} \tag{22a}$$

where $\mathbf{\Xi}_{\perp}$ and $\mathbf{N}_{\perp}$ denote orthogonal compliments of respective matrices.
In I(1) case both from economic, and statistical point of view decomposition (19) is satisfactory. If $\mathbf{Y} \sim I(2)$, among independent cointegrating dependencies, described by the matrix $\mathbf{B}$, may be present both directly stationary CI(2,2), and nonstationary CI(2,1) dependencies. As a consequence, the estimation of $\mathbf{B}$, $\mathbf{A}$, $\mathbf{\Gamma}$ does not suffice to

obtain economically interpretable results (even after structuralisation). In particular it is difficult to interpret cointegration space, defined by vectors concatenating $\mathbf{B}$ matrix. In the model with I(1) variables this space encompass stationary relationships combinations only. In the model with I(2) variables, by definition it is possible to consider among R linearly independent cointegrating relationships both R_0 dependencies making I(2) variables stationary, and R_1 cointegrating regressions CI(2,1). These relationships may be described by matrices $\mathbf{B}_0$ or $\mathbf{B}_1$, $M \times R_0$, $M \times R_1$ respectively. In the space defined by $\mathbf{B}$ lie combinations of both stationary and nonstationary dependencies, hence the interpretation of such combination is impossible in general. Besides parameters estimation of VECM, in the model with I(2) variables it is then necessary to search the obtained from I(1) analysis cointegrating dependencies projection into cointegrating subspace CI(2,2) (it is the space of long- and medium-run equilibrium relationships) and subspace CI(2,1) with more complicated interpretation. These subspaces are mutually orthogonal. Matrix $\mathbf{B}_1$ describes nonstationary relationships. However the vectors concatenating this matrix are cointegrating vectors, because order of integration is decreased. Cointegration CI(2,1) type means (as it was mentioned in section one), that there are similar stochastic trends in variables in the long (strictly: in very long) period, however in the short and medium their paths are not mutually related. Cointegrating dependencies $\mathbf{B}_1\mathbf{Y}_{t-1}$ liquid stochastic trends in first differences, such combinations stay nonstationary in this sense that deviations from such long-run dependencies are generated by the random walk process. Error correction mechanism dominates in the very long run only in the case, when random term from the relationships CI(2,1) is integrated of order one. The essence of cointegration CI(2,1) is that shocks influencing these relationships stop amplifying.
Figure (3) exhibits the projection of R-dimensional "resultant" cointegration space into R_0-dimensional CI(2,2) space and R_1-dimensional CI(2,1) space.
To simplify the Figure three-dimensional cointegration space was regarded, which is CI(2,1) in general, because nonstationary cointegrating dependencies are possible. In the considered case, formulas made possible the three-dimensional space projection into stationary CI(2,2) dependencies plane spanned on independent vectors β_r^0 $(r = 1, 2)$ and into orthogonal with them vector of CI(2,1) dependencies β_1^1 (assuming that $R_0 = 2$, which implies $R_1 = 1$) are necessary.
From the projection into CI(2,1) and CI(2,2) spaces point of view, it is crucial to consider polynomial cointegration relationships, which join these variables (or combinations variables), which did not achieve stationarity by simple cointegrating dependencies (i.e. CI(2,1)). By the definition, CI(2,1) cointegration produce random terms I(1), not cointegrated each over (as random terms from independent and hence mutually orthogonal dependencies). The one possibility to make R_1 combination $\mathbf{B}_1^T\mathbf{Y}_{t-1}$ stationary is to cointegrate them with other I(1) variables, which are potentially present in the model. Such variables are the first increments of integrated of order two noncointegrating dependencies $\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1}$ (vectors concatenating matrix

Figure 3: Cointegration space decomposition in I(2) Model.



$\mathbf{B}_{2\perp}$ do not define relationships which reduced integration order), the number so such increments equals P_2 (by the definition they are I(1)). The necessary condition for model "balance" is then $P_2 = M - R - P_1 = R_1$. Polynomial cointegration relationship is then CI(1,1) regression between levels and first differences of variables (it may be relationship between CI(2,1) combinations of flows categories, and the first differences of mutually noncointegrated stocks). Then polynomial cointegration shall be interpreted as cointegration relationship between flows, while CI(2,2) and CI(2,1) is always (under rather realistic assumption, that no I(3) trends at all) cointegration between stock variables (for example price and money supply).
Three cases are possible. If $P_2 > R$, then model is wrongly specified (in the sense of the economic choice of variables to the system), because not all common trends I(2) will be derived from VECM model by the polynomial co integration.
In the case $P_2 = R$ equilibrium condition arrives, when the one cointegrating relationships in the model are CI(2,1) dependencies. It means that there are no such dependencies between variables, which occur both in the long, and medium period. Number of I(1) trends in such model equals $P_1 = M - 2P_2 = M - 2R$.
If $P_2 < R$, then the following combinations should be determined

$$\mathbf{B}^T\mathbf{Y}_{t-1} - \mathbf{\Lambda}^T\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1} \sim I(0) \tag{22}$$

where $\mathbf{\Lambda}^T = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{\Gamma}\mathbf{B}_{2\perp}(\mathbf{B}_{2\perp}^T\mathbf{B}_{2\perp})^{-1}$ is $R \times M - R - P_1$ matrix (cf. Haldrup 1994, Haldrup 1999).
The element (p_2, r) of matrix $\mathbf{\Lambda}$ may be interpreted as the component of r-th cointegration vector, connected with the p_2-th combination of noncointegrated I(2)

trends first differences.
From the dependency (22) it is clear, that $M - R - P_1 = P_2$ stochastic I(2) trends may be connected by polynomial cointegration with R linearly independent combinations of levels (it derives from the matrix Λ dimensions). Equilibrium condition suggests however, that only R_1 of such combination are linked by polynomial cointegration. Reduced column rank of matrix $\mathbf{\Lambda}$, whereas full row rank (which is the consequence of condition $R \geq P_2$) allows us to remove this apparent contradiction. It should be noted, some combinations concatenating cointegration matrix $\mathbf{B}$ are stationary directly by cointegration CI(2,2). This is the feature of R_0 combinations $\mathbf{B}_0^T\mathbf{Y}_{t-1} \sim I(0)$, which achieve stationarity by simple cointegration. Hence there exist $R - R_0 = R_1$ independent cointegrating relationships (23) only, so the matrix rank $\mathbf{\Lambda}$ equals $R_1 = P_2$ (cf. Haldrup 1999).
Transformation (23) allows us to write (Juselius 2004):

$$\mathbf{B}_1^T\mathbf{Y}_{t-1} - \mathbf{K}^T\Delta\mathbf{Y}_{t-1} \sim I(0) \tag{23}$$

where $\mathbf{K} = \mathbf{B}_{2\perp}\mathbf{\Lambda}\mathbf{\Lambda}^T$, $\mathbf{B}_1 = \mathbf{B}\mathbf{\Lambda}^T$, which after simple transformations lead us to

$$\mathbf{\Lambda}\left(\mathbf{B}^T\mathbf{Y}_{t-1} - \mathbf{\Lambda}^T\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1}\right) \sim I(0) \tag{24}$$

The advantages of equilibrium condition in terms of (24) are as follows. Firstly, it defines relationships of stochastic trends I(2) first differences with cointegration relationships CI(2,1) exclusively, and then direct cointegration dependencies in $\mathbf{B}^T\mathbf{Y}_{t-1}$ are omitted (matrix $\mathbf{K}$ has full column rank contradictory to matrix $\mathbf{\Lambda}$). Secondly, apart from polynomial cointegration matrix $\mathbf{\Lambda}$, more interesting matrix $\mathbf{K}$ is obtained which links first increments variables.
Formula (24) allows us to interpret $\mathbf{\Lambda}$ as projection matrix of "traditional" cointegrating matrix $\mathbf{B}$ into subspace CI(2,1). Simultaneously (Haldrup 1999):

$$\mathbf{B}_0 = \mathbf{B}\mathbf{\Lambda}_\perp^T \tag{25}$$

which means, that $\mathbf{\Lambda}_\perp^T$ is $(R \times R_0)$ projection of matrix $\mathbf{B}$ into subspace of direct dependencies CI(2,2). Problem of cointegrating matrix projections has just solved on the basis of polynomial cointegration. Applying:

$$\mathbf{A}_1\mathbf{K}^T = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T \tag{26}$$

it is possible to make similar projection of adjustments matrices (Juselius 2004).
The CI(2,1) relationships matrix of weights has the form:

$$\mathbf{A}_1 = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T\mathbf{K}\left(\mathbf{K}^T\mathbf{K}\right)^{-1} = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T\bar{\mathbf{K}} \tag{27}$$

Consequently, it holds $\mathbf{A}_0\mathbf{K}_\perp^T = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T$ which allows us to obtain weights of relationships CI(2,2) matrix:

$$\mathbf{A}_0 = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T\mathbf{K}_\perp(\mathbf{K}_\perp^T\mathbf{K}_\perp)^{-1} = \mathbf{A}\bar{\mathbf{A}}^T\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^T\bar{\mathbf{K}}_\perp \tag{28}$$

It means, that $\bar{\mathbf{K}}$ matrix (or its orthogonal compliment) play the same role with respect to the weights matrix as projection matrix $\mathbf{\Lambda}_{\perp}^{T}$ with respect to cointegration space (or its orthogonal compliment). This analogy is however not exact, because the "full" projection matrix has a form $\bar{\mathbf{A}}^{T}\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^{T}\bar{\mathbf{K}}$ or $\bar{\mathbf{A}}^{T}\mathbf{\Gamma}\bar{\mathbf{B}}_{2\perp}\mathbf{B}_{2\perp}^{T}\bar{\mathbf{K}}_{\perp}$ respectively. Recapitulating, in the VECM model with I(2) variables full matrix rank $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$ means the absence of I(2) trends.

The matrices $\mathbf{\Xi}$ and $\mathbf{N}$ are useful not only to decompose into respective subspaces. They are applied to the estimation of VECM model parameters in the case of potential presence of I(2) trends. Apart from the iterative method (Johansen 1994), the most popular is two-stage Johansen procedure (1995b). The name "two-stage Johansen procedure" origins from that, Johansen approach is applied twice. In the first step, $\mathbf{A}$ and $\mathbf{B}$ matrices are estimated (almost identically, as for the model with I(1) variables), ignoring reduced matrix rank $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$. The one difference with respect to traditional I(1) Johansen procedure is that in this step: $\Delta^{2}\mathbf{Y}_{t}=\mathbf{Z}_{0t}$, $\Delta\mathbf{Y}_{t-1}=\mathbf{Z}_{1t}$, $\mathbf{Y}_{t-1}=\mathbf{Z}_{2t}$, whereas all lagged second differences and eventually deterministic terms defines $\mathbf{Z}_{3t}$ variable matrix. The starting model has then more complicated form than the model with I(1) variables):

$$\mathbf{Z}_{0t}=\mathbf{\Gamma}\mathbf{Z}_{1t}+\mathbf{A}\mathbf{B}^{T}\mathbf{Z}_{2t}+\mathbf{\Psi}\mathbf{Z}_{3t}+\Sigma_{t} \tag{29}$$

which means, that the first from residuals regression models (for details of Johansen procedure: Majsterek 1998, Majsterek 2008) will be as follows:

$$\mathbf{R}_{0t}=\mathbf{\Gamma}\mathbf{R}_{1t}+\mathbf{A}\mathbf{B}^{T}\mathbf{R}_{2t}+\tilde{\Sigma}_{t} \tag{30}$$

which requires one more concentrated likelihood function must be additionally constructed, to obtain desired estimates of $\mathbf{\Psi}$, $\mathbf{\Gamma}$, $\mathbf{A}$ and $\mathbf{B}$. Similarly as in the case of one-step procedure, cointegrating matrix $\mathbf{B}$ is obtained from the solution of determinant problem, which due to the Rao (1973) lemma is equivalent with maximization of values "most concentrated" (the simplest from the optimisation point of view) likelihood function value $L_{\max}(\hat{\mathbf{B}})$. The other parameters matrices were defined in the former steps (after successive transformations of explained and explanatory variables matrices) as functions of other matrix parameters. Firstly, estimation of cointegrating matrix allows us to calculate (not: to estimate) the weights matrix $\hat{\mathbf{A}}(\hat{\mathbf{B}})$, secondly weights (adjustments) matrix is useful in the calculation of the medium-run relationships $\mathbf{\Gamma}(\hat{\mathbf{A}},\hat{\mathbf{B}})$ matrix, and finally to obtain the block matrix $\mathbf{\Psi}(\hat{\mathbf{\Gamma}},\hat{\mathbf{A}},\hat{\mathbf{B}})$ consisting of all $\Delta\mathbf{Y}_{t-s}$ $(s=1,\ldots,S-2)$. In all above stages the invariance of FIML estimator is a very important feature contrary to non-invariant estimators based on OLS (for example SUR). This estimation of cointegrating matrix is preceded by construction of successive models and consequently more concentrated (simplified) likelihood function. Initial function has the form $L(\hat{\mathbf{\Psi}},\hat{\mathbf{\Gamma}},\hat{\mathbf{A}},\hat{\mathbf{B}},\hat{\mathbf{\Omega}})$. After defining $\mathbf{\Psi}(\hat{\mathbf{\Gamma}},\hat{\mathbf{A}},\hat{\mathbf{B}})$, VECM model is transformed and its concentrated likelihood function is just $L(\hat{\mathbf{\Gamma}},\hat{\mathbf{A}},\hat{\mathbf{B}},\hat{\mathbf{\Omega}})$. In the next step after defining $\mathbf{\Gamma}(\hat{\mathbf{A}},\hat{\mathbf{B}})$ the next transformation of VECM model is obtained, for which the likelihood function concentrates to $L(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{\Omega}})$, and in the following step - to the determinant problem solution. It should be noted, that estimator of cointegrating matrix is still super-superconsistent. Super-superconsistency of cointegrating matrix $\mathbf{B}$ estimator may be explained, that in the successive steps following concentrated regression (analogously as in the Johansen procedure for I(1)) is obtained:

$$\tilde{\mathbf{R}_{0t}} = \mathbf{A}\mathbf{B}^T\tilde{\mathbf{R}}_{2t} + \tilde{\tilde{\mathbf{\Sigma}}}_t \tag{31}$$

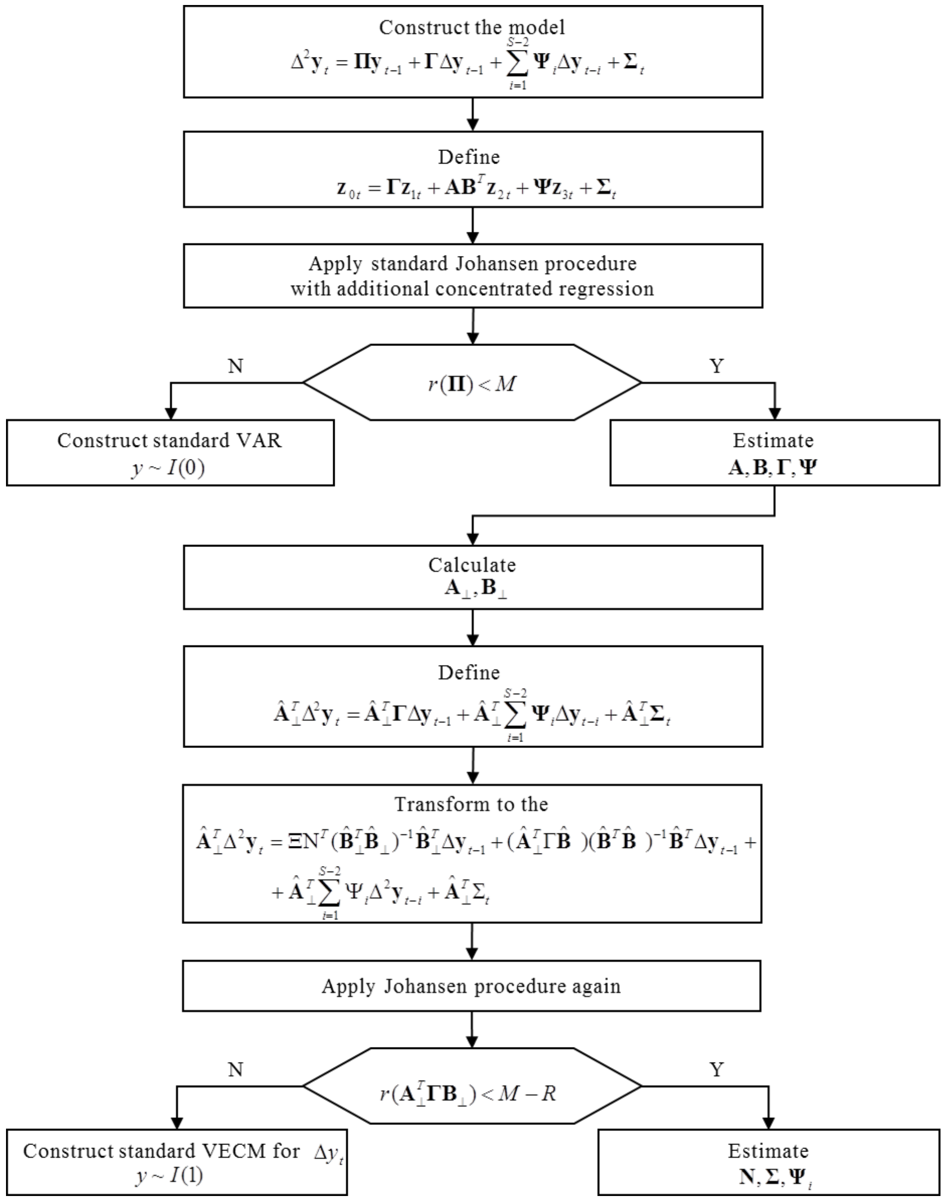
which cointegrate variables integrated of order two $\tilde{\mathbf{R}}_{2t}$ in stationary combinations (both $\tilde{\mathbf{R}}_{0t}$ and $\tilde{\tilde{\mathbf{\Sigma}}}_t$ are stationary). This is because from (31) $\tilde{\mathbf{R}}_{0t}$ are residuals from concentrated regression between $\mathbf{Z}_{0t}$, (stationary second differences) and lagged second differences $\mathbf{Z}_{0,t-s}$. Such residuals are stationary by definition. On the other hand, residuals $\tilde{\mathbf{R}}_{2t}$ origin from the regression between $\mathbf{Z}_{2t}$ (which are integrated of order two levels) and stationary $\mathbf{Z}_{0,t-s}$. Not surprisingly, $\tilde{\mathbf{R}}_{2t}$ are I(2). However it should be emphasised, that super–superconsistency is the feature of all matrix $\mathbf{B}$ estimator, not only the property of measuring CI(2,2) relationships projection $\mathbf{B}_0$ matrix. In particular, the estimates of matrix $\mathbf{B}_1$ (although connected with such relationships, deviations from which are not stationary) are significantly more precise (due to the super - superconsistency) than estimates of defining stationary dependencies the cointegrating matrix in the model with I(1) variables. I(1) analysis invalidity is then caused not by technical reasons, but by the lack of matrix $\mathbf{B}$ and $\mathbf{A}$ interpretability. Additionally it is impossible to project these matrices into respective subspaces. The same disadvantage concerns the orthogonal compliments of these matrices. The two-stage Johansen procedure is then necessary (cf. Figure (4)).
In this step, analogously as in the first stage, the next reduced rank regression problem is solved to obtain $\mathbf{A}_{\perp}^T\mathbf{\Gamma}\mathbf{B}_{\perp}$. It is then useful to premultiply (7) by $\hat{\mathbf{A}}_{\perp}^T$ (this matrix estimate is obtained in the first stage, consequently it is possible to apply estimation methods of the orthogonal compliments matrix for model with I(1) variables). The following modification of VECM is obtained:

$$\hat{\mathbf{A}}_{\perp}^T\Delta^2\mathbf{Y}_t = \hat{\mathbf{A}}_{\perp}^T\mathbf{\Gamma}\Delta\mathbf{Y}_{t-1} + \hat{\mathbf{A}}_{\perp}^T\sum_{i=1}^{S-2}\mathbf{\Psi}_i\Delta^2\mathbf{Y}_{t-i} + \hat{\mathbf{A}}_{\perp}^T\mathbf{\Sigma}_t, \tag{32}$$

because from the orthogonality between $\mathbf{A}$ and $\mathbf{A}_{\perp}^T$, there are not long-run relationship $\mathbf{A}_{\perp}^T\mathbf{A}\mathbf{B}^T\mathbf{Y}_{t-1} = \mathbf{0}$ in the model (32). Instead of M dependencies system comprises $M - R$ relationships for first and second differences. Form (32) is very similar to the model (5), but there are first differences instead of levels, and second differences instead of first ones. Due to relationship (Johansen 1995a, p.135):

$$\hat{\mathbf{B}}_{\perp}\left(\hat{\mathbf{B}}_{\perp}^T\hat{\mathbf{B}}_{\perp}\right)^{-1}\hat{\mathbf{B}}_{\perp}^T + \hat{\mathbf{B}}\left(\hat{\mathbf{B}}^T\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}^T = \mathbf{I}, \tag{33}$$

Figure 4: Two-stage Johansen procedure



model (32) may be transformed to

$$\begin{aligned}\hat{\mathbf{A}}_{\perp}^{T}\Delta^{2}\mathbf{Y}_{t} &= \left(\hat{\mathbf{A}}_{\perp}^{T}\mathbf{\Gamma}\hat{\mathbf{B}}_{\perp}\right)\left(\hat{\mathbf{B}}_{\perp}^{T}\hat{\mathbf{B}}_{\perp}\right)^{-1}\hat{\mathbf{B}}_{\perp}^{T}\Delta\mathbf{Y}_{t-1}+ \\ &\quad+\left(\hat{\mathbf{A}}_{\perp}^{T}\mathbf{\Gamma}\hat{\mathbf{B}}\right)\left(\hat{\mathbf{B}}^{T}\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}^{T}\Delta\mathbf{Y}_{t-1}+ \\ &\quad+\hat{\mathbf{A}}_{\perp}^{T}\sum_{i=1}^{S-2}\mathbf{\Psi}_{i}\Delta^{2}\mathbf{Y}_{t-i}+\hat{\mathbf{A}}_{\perp}^{T}\mathbf{\Sigma}_{t}\end{aligned} \tag{34}$$

and after applying (19) to:

$$\begin{aligned}\hat{\mathbf{A}}_{\perp}^{T}\Delta^{2}\mathbf{Y}_{t} &= \mathbf{\Xi}\mathbf{N}^{T}\left(\hat{\mathbf{B}}_{\perp}^{T}\hat{\mathbf{B}}_{\perp}\right)^{-1}\hat{\mathbf{B}}_{\perp}^{T}\Delta\mathbf{Y}_{t-1}+ \\ &\quad+\left(\hat{\mathbf{A}}_{\perp}^{T}\mathbf{\Gamma}\hat{\mathbf{B}}\right)\left(\hat{\mathbf{B}}^{T}\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}^{T}\Delta\mathbf{Y}_{t-1}+ \\ &\quad+\hat{\mathbf{A}}_{\perp}^{T}\sum_{i=1}^{S-2}\mathbf{\Psi}_{i}\Delta^{2}\mathbf{Y}_{t-i}+\hat{\mathbf{A}}_{\perp}^{T}\mathbf{\Sigma}_{t}\end{aligned} \tag{35}$$

The form (35) is analogous to the starting VECM model (6), and then Johansen method may be then applied again. By the analogy to the first step, in the second one matrices $\mathbf{N}$, $\mathbf{\Xi}$ and $\mathbf{\Psi}$ are estimated, and afterwards mean lag matrix $\mathbf{\Gamma}$ is reestimated, in the second stage of the Johansen approach under reduced rank $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$ condition whereas in the first step this matrix was calculated under full rank of $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$. Consequently, $\mathbf{\Gamma}$ matrix is modified (the very similar is mechanism of modification the form of long-run relationships matrix $\mathbf{\Pi}$ in the case of its full and reduced rank in the I(1) analysis, cf. Majsterek 2008). Just in the second step of Johansen procedure the advantage of decomposition (18) is clear, because due to (18) we are able to obtain $\mathbf{\Xi}$ and $\mathbf{N}$ matrices explicitly. On the basis of (16)-(17) it is not difficult to estimate matrices of parameters $\mathbf{A}_{1\perp}$ and $\mathbf{B}_{1\perp}$ (invariance of FIML estimator is useful again). Similarly, formulas (21)-(22a) allows us to obtain $\mathbf{A}_{2\perp}$ and $\mathbf{B}_{2\perp}$. Projection of cointegrating matrix and weights matrix into subspaces CI(2,2) and CI(2,1) is possible from formulas (27), (28), (25), (23), because all elements of theirs were estimated earlier. Johansen (1995b) and Paruolo (1996) confirmed, that two-stage Johansen procedure is asymptotically the most efficient, equivalent with FIML.
The interesting property of the two-step Johansen method is application of nondirect cointegration phenomenon. It may be noted, that the method explained above simulates gradual equilibrium achievement, and in this sense imitates real adjustment processes in economy. Firstly introductory global long-run equilibrium between economic categories (their increments, so flows) is achieved, in the next step levels (stocks) of these variables are adjusted. In the first step $\mathbf{B}$ and $\mathbf{A}$ matrices are estimated globally (without decomposition into direct and polynomial cointegration), in the second stage the main purpose is decomposition of these cointegrating relationships, and projections of common stochastic trends matrix into I(1) and I(2) subspaces.

There not exists (contrary to I(1) case) separate problem of common stochastic trends I(2) representation parameters estimation. In the model with I(1) variables there exist separation between matrices from "primary" representation (presented in the VECM or VAR model exclusively), and their orthogonal compliments (presented in the common stochastic trends I(1) model only), hence it is necessary to separately estimate parameters of both alternative models (primal and its solution). In the model with I(2) variables, as may be seen for example from (12) such discrimination does not occur. All matrices necessary to estimate parameters of common trends I(2) model were just estimated earlier, in the two-stage Johansen methods.
The problem of cointegration rank determination in the model with I(2) variables is more complicated than in I(1) domain. Two approaches are possible. The simplest is based on sequential procedure. In the first step cointegration rank is determined (by traditional trace test or maximum eigenvalue test) in traditional way, so under assumption about full rank of $\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$. In the next step, actual rank of this latter matrix is found, which allows us to determine not only the number of I(1) and I(2) stochastic trends but matrices ranks of $\mathbf{N}$ and $\mathbf{\Xi}$ too. The classical tests procedure based on likelihood ratio is applied again (trace test is preferred). By the analogy to the finding R cointegrating directions in I(1) analysis, likelihood ratio tests are applied to find the dimensions of $\mathbf{N}$ and $\mathbf{\Xi}$, which again determine dimensions of $\mathbf{B}_0$ and $\mathbf{B}_1$ matrices. Because the cointegration rank test measures the number of stochastic trends for variables levels, number of common trends for first differences (so: I(2) trends)) is a compliment to $M - R$ of $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$ matrix rank. Rank of this matrix is tested the same way as for the matrix rank $\mathbf{\Pi}$ by likelihood ratio test (trace test is suggested by Juselius (1999).
Cointegration rank test (first step) is however correct only, when the hypothesis about full rank of matrix $\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$ is true. If else, Johansen (1995b) and Paruolo (1996) proposed joint test to establish both cointegration rank R, and P_1. Following test statistic Q is applied:

$$Q\left(P_1, R\right) = TRACE\left(R\right) + TRACE\left(P_1/R\right) \tag{36}$$

where $TRACE$ are respective trace statistics values.
The testing starts from the joint hypothesis $H_0 : R = 0 \wedge P_1 = 0$ and in the case of its rejection there are tested: $H_0 : R = 0 \wedge P_1 = 1$, $H_0 : R = 0 \wedge P_1 = 2$,... until, $H_0 : R = 0 \wedge P_1 = M - R - 1$ (equivalent in this step: $H_0 : R = 0 \wedge P_1 = M - 1$). The rejection of all above null hypothesis leads us to verification $H_0 : R = 1 \wedge P_1 = 0$. Procedure terminates, when the null hypothesis may not be rejected. Finally, the determination of P_1 allows us to find P_2, which from polynomial cointegration condition equals R_1. The serious disadvantage of all traditional tests to determine R and P_1, is very low power (Juselius 1999). On the other hand, the advantage is that, similarly as in the model with I(1) variables, it is not necessary to pre-establish deterministic structure (in particular deterministic trend) to properly identify cointegration ranks R and P_1 (Rahbek, Jorgensen, Kongsted 1999). Johansen

(1995b) and Rahbek, Jorgensen, Kongsted 1999 claim, that joint test is consistent with asymptotically correct size.
Recapitulating, consequences of the ignoring the presence of I(2) trends in the model are apparently not serious. The estimator of cointegrating matrix is super-superconsistent (proof in Juselius 2006), and then the estimates of parameters should be, even in not large sample, sufficiently precise. By paradox, in samples with moderate size, will be even more exact than with respect to variables generated by I(1) processes. In the last case the estimators of cointegrating matrix are superconsistent only, which means, that significantly longer sample is required to achieve acceptably high probability of the precise estimates obtaining.
On the other hand, the dimension of matrix $\mathbf{B}$ may be determined wrongly, because cointegration rank test results base on assumption $r(\mathbf{A}_{\perp}^T\mathbf{\Gamma}\mathbf{B}_{\perp}) = M - R$, which is not fulfilled in this case. This may cause, that dimension of common stochastic trends space will be overestimated, so some of cointegrating dependencies will not be identified.
In the multi-equation model with I(2) variables cointegration relationship not necessarily must be long-run (cf. Table 2). Economic implications of medium-run cointegration relationship are almost always connected with stochastic cycles. The defining of P_1 dependencies $\mathbf{B}_{1\perp}^T\Delta\mathbf{Y}_{t-1}$ base on decomposition: $\mathbf{\Gamma} = \mathbf{A}_{1\perp}\mathbf{B}_{1\perp}^T$. The latter come from substitution above decomposition to the formula (18), then:

$$\mathbf{A}_{\perp}^T\Gamma\mathbf{B}_{\perp} = \mathbf{A}_{\perp}^T\left(\mathbf{A}_{1\perp}\mathbf{B}_{1\perp}^T\right)\mathbf{B}_{\perp} \tag{37}$$

However it is worth emphasising, that decomposition $\mathbf{\Gamma} = \mathbf{A}_{1\perp}\mathbf{B}_{1\perp}^T$ is not unique and there may be proposed alternative solutions with respect to the matrix $\mathbf{\Gamma}$.

4 Economic interpretation of I(2) processes

The economic interpretation of both I(1) and I(2) trends and main dependencies in reliable models describes Table 3, which reflects coexistence of deterministic and stochastic trend. All former considerations was conducted implicitly under assumption that no deterministic trend at all (or its existence did not affect considerations results, which was for example in the case of two-stage Johansen method). However it should be considered, how strongly the presence of deterministic tendency (without loss of generality let us assume linear trend) affects the interpretation of I(0), I(1) and I(2) shocks from the Figure (1).
In the Table 3 I(0) shocks were omitted, because interpretation of them stays invariant, irrespective of whether stochastic trends I(2) and/or deterministic trend were reflected in the analysis. I(0) shock, which by definition is not a stochastic trend, has always short-run self-decaying character (cf. Table 4).
The most interesting are stochastic shocks I(1). Their sense differs depending on presumptions with respect to circumstances. In the case of no "dominating" trends, for example in the classical model with I(1) variables, it may be identified I(1) shock

Table 3: Economic interpretation of I(1) and I(2) trends

Type of shock influencing variable	Lack of deterministic trend		Levels of variables are influenced by linear deterministic trend	
	I(2) trend occur	no I(2) trend	I(2) trend occur	no (2) trend
I(1)	medium-run shock (cyclical deviation) or long-run shock persistently disturbing from the long-run tendency	medium-run trend maintained in the long period	medium-run deviation from the long-run cycle	medium-run deviation from the deterministic trend
I(2)	long-run trend	-–	long-run deviations from deterministic trend (similar to long-rung cycle) or long-run shock persistently disturbing from the long-run tendency	—

with stochastic trend, and then I(1) is long-run. It results from the assumed in the models with I(1) variables interpretation that I(0) impulse is short-, whereas I(1) long-run. This is because, that process with the highest integration order (in this case: I(1)) is pushed by stochastic trend, which dominates other shocks. In this way medium-run shock gets reinforce to the long-run and hence it is correct to call it "trend". Deterministic tendency however may theoretically dominate trend I(1) and to "degrade" it to the role of cyclical deviations from deterministic trend. Such inference is proper only under the assumption, that I(1) trend dominates I(0) shocks, but not deterministic trend. It is supposed, that in this direction should be performed the correct economic explanation of the apparent contradiction, if I(1) type nonstationarity occurs in the case of deterministic trend ignoring, but the integration order is zero, when presence of this trend is reflected. If however stochastic trend dominates, which cause permanent disturbance of the variables from its long-run tendency, then it should be expected the identification of stochastic trend I(1), independently whether deterministic trend is considered or not. It should be additionally stressed, that in the case of I(2) trend presence, just this trend dominates I(1) trend independently whether, deterministic trend is statistically significant or not. I(2) trend is always long-run. Similarly as I(1) trend, it may dominate deterministic trend or may be dominated by this trend. In this second case it should be expected the identification of I(2) shocks only, when lack of deterministic tendency is assumed. Integration order of processes generating variables shall not be treated mechanically,

as is done by researchers analysing traditional unit root integration tests results only. Some of variables may cointegrate both with I(1) and I(2) variables (cf. Juselius 2006). Non-unique interpretation of I(1) trends allows us to better understand the economic sense of cointegration CI(2,1). In the model with variables at most I(1) random term I(1) means, that deviations have maintaining character, errors cumulate. In I(2) model, random term I(1) means the introductory, long-run equilibrium (i.e. flows equilibrium between stock categories) achievement in the system.
Supplementary economic interpretation of the integration with different order presents Table 4.

Table 4: Interpretation of variables integration order

Integration order	Interpretation
I(-1)	past changes correction of cumulant
I(0)	shocks are temporary, which means, that changes decay
I(1)	shocks influencing variable are permanent, variables is dominated by its own past, shocks influencing increments are temporary, which means, that changes maintain, but its acceleration does not
I(2)	permanent shocks on increments, acceleration of changes maintain

The problem of economic restrictions imposing and VECM model structuralisation is widely discussed in literature for the case I(1), cf. for example fundamental works: Johansen 1988, Johansen 1994, Juselius 1999, Juselius 2004 and descriptions: Majsterek 2005, Majsterek 2008. It should be noted, that reduced rank of $\mathbf{\Pi}$ assumption is connected with the switch from the model with jointly stationary variables to the model, in which variables integrated of order at least first occur. Such restrictions are imposed on parameters from the VECM model, but not from the common stochastic trends representation. Assuming additionally reduced rank of $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$ matrix, the presence I(2) trends is allowed. In this context I(2) model may be treated as nested case of I(1) model after positive verification of reduced rank $\mathbf{A}_{\perp}^{T}\mathbf{\Gamma}\mathbf{B}_{\perp}$ restrictions. Analogously, I(1) model is treated as nested case of I(0). If however common stochastic trends model is analysed, then the encompassing occur in the opposite direction. In this case I(1) model is nested in I(2), when restriction that $\mathbf{c}|_2$ is the zero matrix is fulfilled. Analogously, when $\mathbf{c}|_1$ is the zero matrix, then I(1) model may be simplified to I(0). However it should be stressed, that if $\mathbf{c}|_2$ is nonzero, then even if $\mathbf{c}|_1$ is zero, we may only conclude lack of medium-run shocks, but not the case of joint stationarity.
More complicated is interpretation of those matrices, which occur both in the dual and primary representation, because restrictions imposed on these parameters have non-unique meaning. There is a danger, that restriction will introduce as supplementary effect such condition, which is from the economic or statistical point of view undesirable. Sense of some such restrictions is easier to understand, comparing simultaneously the interpretation of key economic matrices in the representation (6) and their orthogonal compliments (cf. Table 6).

Table 5: The application of restrictions on matrices in the model with I(2) variables

Restrictions on matrix	Basic Application	Supplementary Application
Π	no joint stationarity restrictions cointegration rank determination	
$\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$	no joint stationarity of first differences restrictions	
Γ	number of medium-run cointegrating relationships determination	
$\mathbf{B}_0$	stationarity of variables combinations test	stationarity tests with respect to the basic variables of the system
$\mathbf{B}_1$	difference stationarity of variables combinations test	difference stationarity tests with respect to the basic variables of the system
$\mathbf{A}_0$	inclusion of economic knowledge concerning adjustment reactions to the relationships CI(2,2) weak exogeneity in CI(2,2) relationships	
$\mathbf{A}_1$	inclusion of economic knowledge concerning adjustment reactions to the relationships CI(2,1) weak exogeneity in CI(2,1) relationships	
$\mathbf{c}\|_1$	stationarity restrictions concerning variables (only after testing restrictions about $\mathbf{c}\|_2$)	
$\mathbf{c}\|_2$	no I(2) trends in variables restrictions	
$\mathbf{B}_{\perp}$	inclusion of economic medium-run theory number of medium-run relationships	
$\mathbf{B}_{2\perp}$	system stationarity testing for first increments of variables quadratic trend removing analysis of I(2) shocks destinations in the system	
$\mathbf{A}_{1\perp}$	medium-run weak exogeneity medium-run adjustments	source of shocks I(1)
$\mathbf{A}_{2\perp}$		source of shocks I(2)
Ψ_i	VECM model order of lags research strict exogeneity in short-run inclusion of short-run economic theory	no impact of chosen lagged variable
Λ	inclusion of knowledge about dynamic equilibrium relationships	
K	inclusion of knowledge about dynamic equilibrium relationships	
$\mathbf{A}_{sim}^2$	short-run weak exogeneity	impact of chosen structural shock exclusion

In some cases, on the basis of matrix $\mathbf{A}_{2\perp}$ elements analysis (or respectively $\mathbf{A}_{1\perp}$) it may be identified the dominant term of baseline stochastic trend (or respectively stochastic cyclical). Such trend may be called autonomous. With respect to autonomous shocks it is sufficient to fulfil the assumption, that these shocks are weakly correlated. If above condition is fulfilled, it is useful to classify shocks as stock or flows, real or nominal and finally demand or supply. However it must be discriminated

Table 6: Interpretation of matrix **A**, **B**, their compliments and projection

Matrix	Interpretation
$\mathbf{A}_0$	Weights of CI(2,2) relationship – parameters connected with stationary deviations from long- and medium-run equilibrium relationship. Deviations "dominate" only in the short-run
$\mathbf{A}_1$	Weights of relationship CI(2,1) – parameters connected with I(1) deviations from long-run equilibrium relationship. Deviations "dominate" both in the short and medium-run
$\mathbf{A}_{1\perp}$	Weights of medium-run relationship – parameters connected with stationary (short-run) deviations from this relationship. Because medium-run cointegration may be identified with cyclical relationships, both with respect to the dominant I(2) trend, as with respect to the long-run dynamic equilibrium state I(0), above deviations are interpreted as short- and medium-run stochastic trends I(1). They are equivalent with the coefficients of stochastic trends I(1). In the long period however these deviations are stationary with respect to I(2) trends (domination effect occur), hence they cumulate to the long-run I(2) trends with respect to the stationary state
$\mathbf{A}_{2\perp}$	Weights of short-run (non-cointegration) relationship – parameters connected with deviations from these relationships (short-, long- and medium-period). These deviations are I(2) with respect to the stationary state, are then equivalent with coefficients connected with the stochastic trends I(2)
$\mathbf{B}_0$	Coefficients of CI(2,2) relationship defining in the short period deviations from the long- and medium-run equilibrium relationships
$\mathbf{B}_1$	Coefficients of CI(2,1) relationship defining in the short and medium period deviations from the long-run equilibrium relationships
$\mathbf{B}_{\perp}$	Coefficients of medium-run relationship (i.e. stochastic cycle relationship). In the short period deviations from this relationship are stationary. Both in the short, as in the medium period these are deviations both from stationary state, and from dominant I(2) trend. In the long-run these deviations are stationary with respect to I(2) trend, however are I(2) with respect to the long-run dynamic equilibrium state
$\mathbf{B}_{2\perp}$	Coefficients of noncointegrating relationship, which acts in the short period. Both in the long and medium period deviations from this relationship are nonstationary, in the long whereas in the long-run their cumulate to the I(2) trend, in the medium - to I(1) trend only.

for example nominal shocks and the shocks influencing nominal category, the latter should be analysed in terms of $\tilde{\mathbf{B}}_{2\perp}$ matrix. Consequently, type of shocks has not many common with the level of its persistence. Contrary to shocks influencing stocks, which by definition are more permanent than shocks influencing flows, stock shocks not necessarily have longer period of impact than flows shocks. Hence the proper identification of economic shocks allows us to identify for example the impact of nominal shocks on the real side of the economy (in the long and medium period) or vice versa. On the other hand the analysis of centripetal (cointegrating) powers in the system is very similar. Very helpful are in this case $\mathbf{A}_0$ and $\mathbf{A}_1$ matrices, which inform us about o adjustments to the relationships CI(2,2) and CI(2,1) respectively. Hence, depending on which variable is governed by adjustment reaction, it may be, by analogy to the centrifugal shocks, regarded nominal or real categories adjustments, demand or supply adjustments and flows or stocks adjustments. Depending on, whether these are error correction reactions with respect to CI(2,2) or CI(2,1) it may be classified long- and medium-run adjustment or long-run adjustment only.

Duality in the interpretation of matrix $\mathbf{A}_{1\perp}$origins from, that depending on assumed perspective (cf. Figure 1) medium-run cointegration relationship might be treated as centrifugal (common trend I(1)) or centripetal (stationary across stochastic cycle).
Analogously it may be considered the matrix $\mathbf{B}$ and its orthogonal compliments interpretation. The rows of matrix $\mathbf{B}_{2\perp}$ identify variables, which are the most sensitive on double cumulative long-run shocks in the system. With respect to the matrix $\mathbf{B}_{1\perp}$ similar interpretation is impossible. In the case of the analysis of column in the Table 6 it should be noted, that the columns of orthogonal compliments cover coefficients of defining random term dependencies, whereas columns of several $\mathbf{A}$ projection matrices and their orthogonal compliments cover parameters connected with (stationary or not) random factors. The most complicated and the less discussed in the literature is a matrix $\mathbf{B}_{1\perp}$ interpretation. Its elements may be defined as coefficients of medium-run dependencies.

5 Whether I(3) model has economic explanation?

Model (6) may be transformed to the form:

$$\Delta^3\mathbf{Y}_t = \mathbf{\Pi}\mathbf{Y}_{t-1} + \mathbf{\Gamma}\Delta\mathbf{Y}_{t-1} + \mathbf{\Phi}\Delta^2\mathbf{Y}_{t-1} + \sum_{s=1}^{S-3}\mathbf{\Psi}_{s+1}\Delta^3\mathbf{Y}_{t-s} + \mathbf{\Sigma}_t \tag{38}$$

where $\mathbf{\Phi} = \left(\sum_{s=1}^{S-2}\mathbf{\Psi}_s - \mathbf{I}\right)$.
The necessary condition of no I(3) trends in the model is (proof in Johansen 1995a) that $M \times M$ matrix:

$$\mathbf{M} = \mathbf{\Gamma}\mathbf{B}\left(\mathbf{B}^T\mathbf{B}\right)^{-1}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{\Gamma} - \mathbf{\Phi} \tag{39}$$

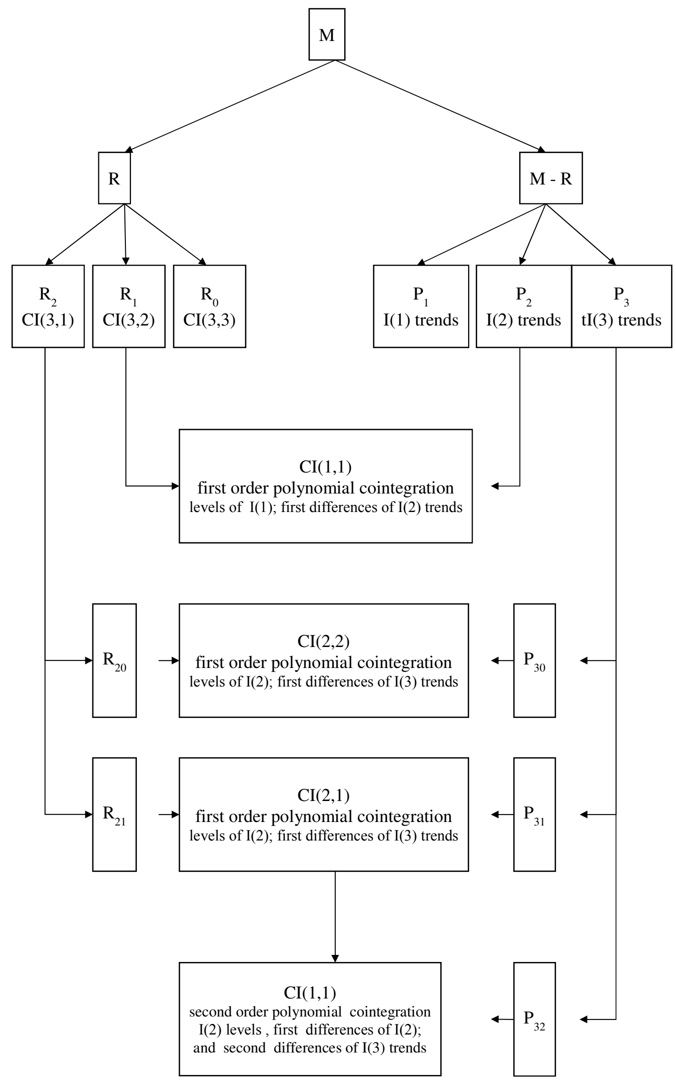
has full column rank. Such matrix may not be decomposed. It means, that similarly as full matrix rank of $\mathbf{\Gamma}$ guarantees for all m ($m = 1, \ldots, M$) stationarity of first differences - full matrix rank $\mathbf{M}$ assures for all m stationarity of second differences. It is worth emphasising, that only reduced rank of $\mathbf{\Gamma}$ and full rank of $\mathbf{M}$ jointly constitute the correct I(2) condition.
It is clear, that the most of the generating variables processes may be transformed to stationary by at most application of difference filter twice or is not integrated at all. On the other hand, especially in the less stable economies, it may be possible, that some variables may be, particularly in the economic disturbances periods, integrated of order three (the possibility of such processes reflect Juselius 2004 and Burke and Hunter 2005, p.159). It is supposed, that such variable may be prices in terms of permanent, nonstable hyperinflation (Juselius 2004). Such type of hyperinflation may be identified with inflationary process, which increase is generated by the random walk, so inflation is I(2), whereas prices are I(3). The main hindrance in such models with I(3) variables application is that, such phenomena, as hyperinflation

are rather not permanent. Only disposing the data with sufficiently high frequency, such processes may be modelled. It narrows the potential application of model with I(3) variables to nonstable state economies (in the case of Poland these techniques might be useful for economy historicians analysing for example hyperinflation after first world war). The next problem is that nontypical variables (i.e. I(3)) do not cointegrate with any basic economic categories. This suggests that hyperinflation may be considered as for example lack of long-run equilibrium between wages (integrated of order at most two) and prices (I(3)). In the case of I(3) variables the analysis of model (39) should be applied, simultaneously in the I(3) domain the matrix $\tilde{\mathbf{M}} = \mathbf{A}_{2\perp}^T(\mathbf{\Gamma B}(\mathbf{B}^T\mathbf{B})^{-1}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{\Gamma} - \mathbf{\Phi})\mathbf{B}_2$ has no full rank $(M-R-P_1)$, and may be decomposed as follows:

$$\tilde{\mathbf{M}} = \mathbf{T}\mathbf{H}^T \tag{40}$$

where $\mathbf{T}$ and $\mathbf{H}$ matrices are $(M-R-P_1) \times P_2$.
Contrary to the models with I(2) variables: identity $(M-R-P_1) = P_2$ does not occur, because P_3 stochastic trends I(3) additionally occur.
Apart from classical long-run dependencies $\mathbf{\Pi Y}_{t-1}$, and short-run relationships $\sum_{i=1}^{I-3} \mathbf{\Psi}_{i+1}\Delta^3\mathbf{Y}_{t-i}$ there are two types of medium-run relationship (in the Table 7 denoted as (a) and (b) respectively): with first $\mathbf{\Gamma}\Delta\mathbf{Y}_{t-1}$ and second differences $\mathbf{\Phi}\Delta^2\mathbf{Y}_{t-1}$. Their potential interpretation is connected with cyclical deviations with longer period of fluctuations (first increments) and overlapping deviations, for example seasonal (second differences). Long-run relationships between the variables integrated of order three have three forms: direct long - and medium-run equilibrium relationship CI(3,3) and nonstationary cointegrating relationships both CI(3,2), and CI(3,1). Cointegration matrix and weights matrix are decomposed (by reliable projections) on three parts: $\mathbf{B} = \left[\mathbf{B}_0 \vdots \mathbf{B}_1 \vdots \mathbf{B}_2\right]$ and $\mathbf{A} = \left[\mathbf{A}_0 \vdots \mathbf{A}_1 \vdots \mathbf{A}_2\right]$ respectively, where $\mathbf{B}_0$ is $M \times R_0$ matrix of CI(3,3) dependencies, $\mathbf{B}_1$ - $M \times R_1$ matrix of CI(3,2) relationships, whereas $\mathbf{B}_2$ $M \times R_2$ matrix explains CI(3,1) relationships. Analogously the components of weights matrix are defined. All matrices projections are mutually orthogonal. Similar relationships take place for the orthogonal compliments: $\mathbf{B}_\perp = \left[\mathbf{B}_{1\perp} \vdots \mathbf{B}_{2\perp} \vdots \mathbf{B}_{3\perp}\right]$ and $\mathbf{A}_\perp = \left[\mathbf{A}_{1\perp} \vdots \mathbf{A}_{2\perp} \vdots \mathbf{A}_{3\perp}\right]$, whereas the latter define common trends: I(1), I(2) and I(3) respectively. Dependencies between first differences are defined by stationary combinations $\mathbf{B}_{1\perp}^T\Delta\mathbf{Y}_{t-1}$ (the number of which is P_1) and integrated of order one relationships $\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1}$, which number equals P_2. Relationships $\mathbf{B}_{3\perp}^T\Delta\mathbf{Y}_{t-1}$ are not cointegrating, and then define combinations integrated of order two. The links between second differences consist of stationary dependencies $\mathbf{B}_{2\perp}^T\Delta^2\mathbf{Y}_{t-1}$ and integrated of order one noncointegrating combination: $\mathbf{B}_{3\perp}^T\Delta^2\mathbf{Y}_{t-1}$. Strictly short-run dependencies are described by relationships between third increments.
There may be present two types of dynamic cointegrating dependencies between levels and first differences, however only the combinations $\mathbf{B}_1^T Y_{t-1}$ and $\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1}$

Figure 5: Model for I(3) variables



are stationary CI(1,1). There additionally exist relationships between $\mathbf{B}_2^T\mathbf{Y}_{t-1}$ and $\mathbf{B}_{3\perp}^T\Delta\mathbf{Y}_{t-1}$. Some of them are stationary CI(2,2), but some are CI(2,1). To make these dynamic dependencies stationary, these combinations should be cointegrated with integrated of order one combinations $\mathbf{B}_{3\perp}^T\Delta^2\mathbf{Y}_{t-1}$. Consequently, such type of polynomial cointegration, although stationary, has more complicated character. It links the combinations of levels and first differences with the second differences. It may be proposed for such polynomial cointegration relationships the name complex polynomial cointegration (second order), contrary to relationships between levels and first differences, which may be called as simple (first order). Second order dynamic cointegration dependencies may be for example the relations between price acceleration (in terms of hyperinflation) and variable being a combination of integrated of order one categories and the increase of (stock) variables integrated of order two. Besides the classical I(2) polynomial cointegration condition $R_1 = P_2$ (in the I(3) model this condition may be renamed to first order stationary polynomial cointegration condition) there shall be reflected first order polynomial nonstationary cointegration condition: $R_2 = P_3$. This condition should be however decomposed. Among the first order polynomial cointegration relationships, relating I(2) variables,

Table 7: Cointegrating relationships in the model for I(3) variables

Type of relation-ship	Static relationships			Polynomial relationships (dynamic)	
	Long-run relationship	Medium-run relationships (a)	Medium-run relationships (b)	simple	complex
CI(3,3)	R_0 dependencies $\mathbf{B_0^T} Y_{t-1}$				
CI(3,2)	R_1 dependencies $\mathbf{B_1^T} Y_{t-1}$				
CI(3,1)	R_1 dependencies $\mathbf{B_2^T} Y_{t-1}$				
CI(2,2)		P_1 dependencies $\mathbf{B}_{1\perp}^T \Delta \mathbf{Y}_{t-1}$		relationship $\mathbf{B_2^T} Y_{t-1}$ and $\mathbf{B}_{3\perp}^T \Delta \mathbf{Y}_{t-1}$	
CI(2,1)		P_2 dependencies $\mathbf{B}_{2\perp}^T \Delta \mathbf{Y}_{t-1}$		relationship $\mathbf{B_2^T} Y_{t-1}$ and $\mathbf{B}_{3\perp}^T \Delta \mathbf{Y}_{t-1}$ $(\mathbf{w}_{1,t-1})$	
CI(1,1)			M_1 dependencies $\mathbf{B}_{2\perp}^T \Delta^2 \mathbf{Y}_{t-1}$	relationship $\mathbf{B_1^T} Y_{t-1}$ and $\mathbf{B}_{2\perp}^T \Delta \mathbf{Y}_{t-1}$	relationship $(\mathbf{w}_{1,t-1})$ and $\mathbf{B}_{3\perp}^T \Delta^2 \mathbf{Y}_{t-1}$

there are both relationships CI(2,2), which number equals R_{20} and relationships CI(2,1), which number is R_{21}. Hence, it must be fulfilled: $R_{20} = P_{30}$ and $R_{21} = P_{31}$, additionally $P_{30} + P_{31} = P_3$. Second order polynomial cointegration links integrated of order one combination made from CI(2,1) dependencies and second differences of I(3) variables, i.e. $R_{21} = P_{32}$. From the above condition it results rather serious restriction. The number of second differences of I(3) trends is P_3, exactly the same is the number of second differences. Indeed, second order polynomial cointegration condition should be $R_{21} = P_3$, which implies $R_{20} = 0$, to avoid contradiction with first order cointegration condition. The latter means that first order polynomial cointegration relationship must not be CI(2,2). The new element of the model with

I(3) variables compare to model with I(2) variables is that stochastic trends I(3) may in two ways participate in polynomial cointegration dependencies (first and second differences of such trends separately). I(2) trends may be present in the polynomial cointegration dependencies as first differences only (the second one stationary by definition and should not cointegrate at all), I(1) trends are not linked by polynomial cointegration dependencies at all.
It should be also noted, that $\mathbf{B}_{2\perp}^{T}$ matrices, present in relationships $\mathbf{B}_{2\perp}^{T}\Delta\mathbf{Y}_{t-1}$ and $\mathbf{B}_{2\perp}^{T}\Delta^{2}\mathbf{Y}_{t-1}$, are probably not the same matrices (as in beginnings of I(2) analysis it were not be discriminated $\mathbf{B}_{1}^{T}$ and $\mathbf{B}_{1\perp}^{T}$ matrices, cf. for example Johansen 1994; Paruolo 2000). It may be then possible, that space I(3) dimension is greater than I(2) space, relationship $\mathbf{B}_{2\perp\perp}^{T}\Delta^{2}\mathbf{Y}_{t-1}$ must be regarded, where matrix $\mathbf{B}_{2\perp\perp}^{T}$ will be orthogonal both with respect to $\mathbf{B}_{2\perp}^{T}$, and cointegrating matrix $\mathbf{B}_{2}^{T}$.

6 Models comparison

Comparison of models with I(0) variables, I(1), I(2) presents Table 8. It must not necessarily be identified the model (4) with the case I(0), the model (5) with the case I(1), and model (6) with the case I(2). All the mentioned above representations are indeed the same model, only isomorphically transformed. Then, it is not wrong, for example, using the model (6) for case I(1), to test possibility of the stochastic trends I(2) exclusion. The application of the model (4), when common stochastic trends occur is still not wrong, but is uncomfortable from the cointegrating matrix and weights matrix estimation point of view. All of representations (4), (5), (6) have the same solution (dual representation), which is dependent on the highest integration order of variables in the model.
Order of lags of the VAR model has the substantial meaning not only in the context of correct model dynamic structure, degrees of freedom number, but also integration order. In the multi-dimensional case, for the static model it follows:

$$\Delta\mathbf{Y}_{t} = -\mathbf{Y}_{t-1} + \mathbf{\Sigma}_{t} \tag{41}$$

which lead us to the identical inferences. The matrix $\mathbf{\Pi} = -\mathbf{I}$ has by definition full rank, which implies joint stationarity. It origins from that lags order of VAR model should be at least one level higher than expected integration order of variables used in the model. VAR (0) model in the form of (6) is under above assumptions:

$$\Delta^{2}\mathbf{Y}_{t} = -\mathbf{Y}_{t-1} - \Delta\mathbf{Y}_{t-1} + \mathbf{\Sigma}_{t} \tag{42}$$

which also confirms inference about VAR(0) stationarity. Consequently, VAR model (1) in the form (6) is:

$$\Delta^{2}\mathbf{Y}_{t} = (\mathbf{\Pi}_{1} - \mathbf{I})\,\mathbf{Y}_{t-1} - \Delta\mathbf{Y}_{t-1} + \mathbf{\Sigma}_{t} \tag{43}$$

which means, that $r(-\mathbf{A}_{\perp}^{T}\mathbf{B}_{\perp}) = r(\mathbf{I}_{T\times(M-R)}) = M - R$, so there are no I(2) trends in the model. Dependency condition formulation between lags order of VAR model

Table 8: Comparison of the VAR representations for different types of models

I(0) Variables	I(1) Variables	I(2) Variables
VAR model for levels variables recommended	VECM model (5) recommended (levels and first increments variables)	model (6) recommended (levels and first two differences variables)
full rank of Π	reduced rank of Π, decomposition of Π	reduced rank of Π, decomposition of Π
solution in the form of MA representation	solution in the form of I(1) trends representation	solution in the form of I(2) trends representation
zero rank of C	nonzero rank of C, decomposition of C, zero rank of C_2	nonzero rank of C_2, decomposition of C_2, zero rank of C_3
	full rank of $\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$	reduced rank of $\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$, decomposition of $\mathbf{A}_{\perp}^{T}\Gamma\mathbf{B}_{\perp}$
consistent estimator	superconsistent estimator	super-superconsistent estimator
short-run analysis	short-and long-run analysis	short-, long- and medium-run analysis
SUR estimation	Johansen method	two-step Johansen method iterative FIML
non-cointegrating relationships CI(0,0)	direct cointegration between levels of variables exclusively	direct (static) and polynomial (dynamic) cointegration
short-run relationships maintained in the long	long-run cointegration relationships	long- and medium-run cointegration relationships
random term is stationary by definition	cointegrating relationships with stationary random term	both stationary and I(1) cointegrating relationships

and integration order in the form of strong inequality $S > d$ is caused not by technical reasons (weak inequality would be sufficient from this criterion), but by the fact, that in the case $S = d$ it must not be estimated the short-run relationships irrespectively of whether we apply the model (5) or (6). In this context it may be augmented the Table 5 by additional meanings. Restriction $\mathbf{\Gamma}_1 = \mathbf{0}$ in the VECM model (assuming that higher lags were not reflected) is equivalent with joint stationarity assumption with respect to first differences variables used in the model, whereas restriction $\mathbf{\Pi}_1 = \mathbf{0}$ in the VECM model (assuming that there are no higher lags) is connected with assumption concerning joint stationary levels of variables.

In the model with I(2) variables the number of the presented here relationships is unproportionally more than in I(1) case. Table 9 illustrates this problem.

In the model with I(2) variables there are many types of cointegration due to different criteria. From the integration order point of view, there are both CI(2,2), CI(2,1), CI(1,1) dependencies. There are both relationships making directly stationary random term, and nonstationary cointegration, similarly there are both traditional long-run and medium-run cointegration relationships. Finally non-standard polynomial cointegration is present: relationship between levels and first differences, which must not be easily classified as long- or medium-run. Additionally, there are non-cointegrating, both stationary (short-run), and I(1) dependencies.

Table 9: Relationships in the the I(2) and I(1) model - comparison

<table>
<tr><td colspan="5">I(2) Model</td></tr>
<tr><td rowspan="2">Integration order of combination</td><td colspan="3">Simple relationships</td><td rowspan="2">Complex relationship</td></tr>
<tr><td>Long-run</td><td>Medium-run</td><td>Short-run</td></tr>
<tr><td rowspan="2">I(0)</td><td colspan="2">R_0 CI(2,2) relationships $\mathbf{B}_0^T\mathbf{Y}_{t-1}$</td><td rowspan="2">$M$ $\Psi_i\Delta^2\mathbf{Y}_{t-i}$ relationships</td><td rowspan="2">$R_1 = P_2$ polynomial cointegration CI(1,1) relationships</td></tr>
<tr><td></td><td>P_1 CI(1,1) relationships $\mathbf{B}_{1\perp}^T\Delta\mathbf{Y}_{t-1}$</td></tr>
<tr><td>I(1)</td><td>R_1 CI(2,1) relationships $\mathbf{B}_1^T\mathbf{Y}_{t-1}$; P_1 common trends I(1)</td><td>P_2 non-cointegration relationships $\mathbf{B}_{2\perp}^T\Delta\mathbf{Y}_{t-1}$</td><td></td><td></td></tr>
<tr><td>I(2)</td><td>P_2 common trends I(1)</td><td></td><td></td><td></td></tr>
<tr><td colspan="5">I(1) Model</td></tr>
<tr><td rowspan="2">Integration order of combination</td><td colspan="3">Simple relationships</td><td rowspan="2">Complex relationship</td></tr>
<tr><td>Long-run</td><td>Medium-run</td><td>Short-run</td></tr>
<tr><td>I(0)</td><td colspan="2">R CI(1,1) relationships $\mathbf{B}^T\mathbf{Y}_{t-1}$</td><td>$M$ $\Psi_i\Delta\mathbf{Y}_{t-i}$ relationships</td><td>no such relationships</td></tr>
</table>

To clarify the order of complexity of I(2) model it shall be compared the part of the Table 9 connected with I(2) model with this part, which presents relationships in the model with I(1) variables.

7 Empirical example

Let us consider the model constructed by Majsterek and Kelm (2007). The relationships linking wage and prices were conducted. After performing I(1) VECM analysis the initial system was modified.

The modified I(1) system comprised monetary aggregate M2 (lm2p), consumer price index (lcpi), GDP (lgdp), proxy of the BS effect (lbs), unit labour costs (lulc) as well as weakly exogenous nominal interest rate (RO) and real exchange rate (lrer). The weak exogeneity tests strongly supported conditioning the system on the real exchange rate and, potentially, on the proxy of the Balassa-Samuelson effect (wider discussion: Majsterek and Kelm 2007), then the number of variables in the system $M = 5$. The trace test results suggested three-dimensional cointegration space.

The I(2) analysis was conducted within the two-step Johansen procedure. The joint cointegration rank was tested by means of the Paruolo (1996) test. The number of cointegrating vectors R and the number of I(1) stochastic trends P_1 was identified jointly by means of the Q statistic:

$$Q\left(P_1, R\right) = TRACE\left(R\right) + TRACE\left(P_1/R\right) \tag{44}$$

where $TRACE$ - values of respective trace statistic.
The results of the cointegration rank test are summarized in Table 10. The outcomes confirmed the conclusions derived from the I(1) models: it was justified to consider the cumulated unit roots in the data generating processes as the Q test suggests the presence of two both long and medium term C(2,2) cointegrating vectors (direct, stationary cointegration) and one non-direct cointegrating relationship C(2,1) which is nonstationary in the considered sample but which becomes stationary in the long-term.

Table 10: The joint cointegration test for the I(2) model (critical values in brackets)

M-R	R	Q(P1,R)					
5	0	646.04 [191.9]	454.36 [161.9]	333.98 [137.0]	256.87 [114.9]	205.59 [96.5]	194.34 [82.6]
4	1		452.64 [132.0]	264.46 [107.9]	189.19 [87.9]	118.32 [71.3]	111.92 [59.0]
3	2			337.79 [82.3]	149.63 [64.2]	77.85 [49.7]	66.10 [39.3]
2	3				113.28 [44.5]	31.02* [31.6]	22.97 [23.0]
1	4					16.85 [17.6]	5.02 [10.6]
P_2	5	5	4	3	2	1	0

*) – denotes chosen null hypothesis

According to the Johansen procedure the estimation of the cointegrating vectors and adjustment matrices was performed taking $R = 3$ in the first step and then assuming $P_2 = 1$ i.e. the presence of one double unit root in DGP. It is then possible to identify the sources of the long I(2) shocks in the model as well as the variables affected by these shocks. Additionally, basing on the cointegrating vectors renormalisation and the analysis of the adjustment parameters one can try to find a link between the estimated cointegrating vectors and theory-based models.

Table 11: The estimates of orthogonal compliments for the I(2) model

Matrix rows	lm2p	lcpi	lgdp	lbs	Lulc
$\mathbf{A}_{2\perp}$	0.1202	-0.4948	0.0381	-0.0670	0.0280
$\mathbf{B}_{2\perp}$	-0.2770	-0.1168	-0.0760	-0.0404	-0.1167
$\mathbf{A}_{1\perp}$	-0.7413	-0.2901	-2.4188	0.9549	3.6279

The main economic inference from the Table 11 is as follows. Price shocks rather than money demand shocks produce the I(2) behaviours (elements of $\mathbf{A}_{2\perp}$) in the

system (the influence of the monetary shocks is about four times weaker, 0.12), whereas money demand is I(2) variable in the sense that it import such shocks (elements of $\mathbf{B}_{2\perp}$). Then this is additional argument to recognise famous Fisher equation as the money demand rather than price equation. Moreover, as the compared parameters (connected with money and prices) have opposite signs the results may be perceived as a justification of the hypothesis according to which disinflationary process strongly outperformed the inflationary pressure connected with the monetary expansion. Secondly, there are three variables that accumulate I(2) shocks (cf. elements of $\mathbf{B}_{2\perp}$): the nominal monetary aggregate (-0.28), the consumer prices (-0.12) and the unit labour costs (-0.12). The elements of $\mathbf{A}_{1\perp}$ matrix inform us about the sources of medium – run stochastic shocks (stochastic cyclical).
The results of the "stationary" cointegrating vectors concatenating $\mathbf{B}_0$ analysis are as follows. One of the possible renormalizations leads to the following long- and medium-run CI(2,2) relationships

$$\begin{aligned} lm2p &= 2.066lgdp + 1.081lulc - 0.122lcpi + 0.195lbs - 0.450lrer - 14.6RO \\ lbs &= 0.233lm2p - 0.306lgdp + 0.050lulc - 0.059lcpi + 0.043lrer + 2.62RO \end{aligned} \tag{45}$$

The first cointegrating vector may be interpreted in terms of the money demand function. The scale variable is represented by GDP with the elasticity exceeding 2. The estimates of the prices' elasticities seem to be a little bit confusing as the CPI elasticity is negative, whereas the ULC elasticity exceeds unit. Such a result might be acceptable however as it may be connected with the overlapping effect between CPI and ULC (the estimates sum up to unity). A speculative demand for money is represented by the real exchange rate and nominal interest rate. The long-term parameters connected with these two variables are negative and confirm the presence of the mechanism responsible for the decrease of the money demand in the case of a zloty depreciation and in the case of an increase of the return rates from the monetary substitutes.
The interpretation of the second C(2,2) cointegrating vector is less obvious. Let us recall that Balassa-Samuelson is approximated by the ratio of the wages in sheltered and open sectors. In such a case the parameters in the second cointegrating vector allow us for an identification of the main causes of the differences in the dynamics of the wages in both sectors. If so, the monetary expansion should be interpreted as the important reason of the increase of the BS ratio whereas the higher dynamics of GDP leads to the faster wage growth in the open sector. The general conclusion that may be derived from such results is clear: the wages in the production sector are tied to the labour productivity whereas the increase of the wages in the service sector is affected by the indexation mechanisms.
The intepretation of the one vector of $\mathbf{B}_0$ is more complicated and requires performing polynomial cointegration investigations. Hence, this part of analysis may be omitted. The results confirmed, there are strong I(2) symptoms in the system if we focus on the transition period 1995-2005. The empirical analysis of the relationship between prices

and wages calls for the application of the I(2) cointegration techniques. Secondly, the I(2) properties of the nominal sphere of the Polish economy should be perceived as a consequence of the strong disinflation process but not as a result of the tight monetary policy preventing form the excessive monetary expansion.

8 Conclusions

Models with I(2) variables indeed existed in the past, but from different reasons this fact was ignored. It was cause both by the shortage of statistic and economic knowledge, lack of proper software (in the past), and many either reasons. Among main reasons of the I(2) analysis omission there were difficulties connected with complicated interpretation, especially difficulty with finding the sense of nonstationary cointegration CI(2,1), which violated general opinions (maintained by too literally treated Granger Representation Theorem), that the economic sense has only such long-run relationship, from which deviations are stationary. In the paper it was proved, that cointegration with random term I(1) has interesting economic interpretation. It was also presented, that statistic and economic consequences analysis I(2) ignoring, in the case when such processes really exist in the analysed system can (but not necessarily) be serious. These conditions serve to simplify I(2) analysis towards an easy and popular cointegration with I(1) variables.
The mentioned in the paper difficulties connected analysis I(2) application must not limit its arising popularity. From the earlier considerations it may be derived, that there exist statistic and economic reasons of its application.

Acknowledgements

The paper is based on the chapter 3 from the author's monograph "Wielowymiarowa analiza kointegracyjna w ekonomii" WUŁ 2008, although some problems were omitted, on the other hand a few questions are enlarged.

References

[1] Banerjee A., Dolado J. J., Galbraith J. W., Hendry D. F. (1993), *Co-integration, Error Correction and the Econometric Analysis of Non-stationary Data*, Oxford University Press, Oxford.

[2] Burke S., Hunter J., (2005), *Modelling Non-Stationary Economic Time Series: A Multivariate Approach*, Palgrave Macmillan, Houndmils, Basingstoke.

[3] Dickey D. A., Pantula S. G. (1987), Determining the Order of Differencing in Autoregressive Processes, *Journal of Business and Economic Statistics* 5, 455–462.

[4] Diebold F., Nerlove M. (1990), Unit Roots in Economic Time Series: A Selective Survey, [in:] *Advances in Econometrics. Co-Integration, Spurious Regressions and Unit Roots 8*, [eds.:] T. Thomby, G. F. Rhodes jr., JAI Press Inc, London, 3–70.

[5] Engle R. F., Granger C. W. J. (1987), Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica* 55, 251–276.

[6] Ericsson N. R., Campos J., Hong-Anh Tran (1991), PC-GIVE and David Hendry's Econometrics Methodology, *International Finance Discussion Papers* 406.

[7] Flores R., Szafarz A. (1996), An Enlarged Definition of Cointegration, *Economics Letters* 50, 193-195.

[8] Granger C. W. J., Newbold P. (1974), Spurious Regressions in Econometrics, *Journal of Econometrics* 2, 111-120.

[9] Haldrup N. (1994), The Asymptotics of Single–Equation Cointegration Regressions with I(1) and I(2) variables, *Journal of Econometrics* 63, 153-181.

[10] Haldrup N. (1999), An Econometric Analysis of I(2) Variables, [in:] *Practical Issues in Cointegration Analysis*, M. McAleer, L. Oxley, Blackwell Publishers, Oxford, 179–234.

[11] Hendry D. F., Pagan A. R., Sargan J. D. (1984), Dynamic Specification, Chapter 18, [in:] *Handbook of Econometrics*, [eds.;] Z. Griliches, M. D. Intriligator, North Holland, Amsterdam, 1023–1100.

[12] Johansen S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control* 12, 231-254.

[13] Johansen S. (1994), *An Algorithm for Estimating the Cointegrating Relations in Vector Autoregressive Process Allowing for I(2) Variables*, mimeo.

[14] Johansen S. (1995a), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

[15] Johansen S. (1995b), A Statistical Analysis for I(2) Variables, *Econometric Theory* 11, 25-29.

[16] Johansen S. (1997), A Likelihood Analysis of the I(2) Model, *Scandinavian Journal of Statistics* 24, 433-462.

[17] Juselius K. (1999), Price Convergence in the Long Run and Medium Run. An I(2) Analysis of Six Price Indices, [in:] *Cointegration, Causality and Forecasting, Festschrift in Honour of Clive W.J. Granger*, [eds.:] R. Engle, H. White, Oxford University Press, Oxford, 301–325.

[18] Juselius K. (2004), Inflation, Money Growth and the I(2) Analysis, [in:] *New Directions in Macromodelling*, [ed.:] A. Welfe, Elsevier, Amsterdam, 69–106.

[19] Juselius K. (2006), *Cointegrated VAR Model. Methodology and Applications*, Oxford University Press, Oxford.

[20] Kufel T. (2001), "Nonsense Correlation between Time Series" -– historia badań symulacyjnych procesów zintegrowanych, [in:] *Dynamiczne modele ekonometryczne, Materiały zgłoszone na VII Ogólnopolskie Seminarium Naukowe 4–6 września 2001*, Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń, 119–128.

[21] Majsterek M. (1998), Zastosowanie procedury Johansena do analizy sprzężenia inflacyjnego w gospodarce polskiej, *Przegląd Statystyczny* 45/1, 113–130.

[22] Majsterek (2005), Restrykcje w analizie kointegracyjnej, *Przegląd Statystyczny* 52/1, 55–74.

[23] Majsterek M. (2008), *Wielowymiarowa analiza kointegracyjna w ekonomii*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2008.

[24] Majsterek M., Kelm R. (2007), Relationship Between Wages and Prices in the Polish Economy. An I(2) Approach, [in:] *MACROMODELS'2006*, [eds.:] W. Welfe, A. Welfe, ABSOLWENT, Łódź, 45–64.

[25] Osiewalski J., Pipień M. (1999), O stacjonarności procesów ARMA, *Przegląd Statystyczny* 46/1, 43-50.

[26] Paruolo P. (1996), On the Determination of Integration Indices in I(2) Systems, *Journal of Econometrics* 72/1-2, 313–356.

[27] Paruolo P. (2000), A Characterization of I(2) VAR(1) Process, *Econometric Theory* 17, 235–258.

[28] Phillips P. C. B. (1991), Optimal Inference in Cointegrated Systems, *Econometrica* 59, 283–306.

[29] Rahbek A., Jorgensen C., Kongsted H. Ch. (1999), Trend-Stationarity in the I(2) Cointegration Model, *Journal of Econometrics* 90, 265–289.

[30] Rao C. R. (1973), *Linear Statistical Inference and its Applications*, John Wiley and Sons, New York.

[31] Romański J., Strzała K. (1995), Testing for Integration and Cointegration with Macroeconomic Indicators for Poland, [in:] *Proceedings of the "Macromodels'1991" Conference*, [eds.:] W. Welfe, W. Florczak, Absolwent, Łódź, 263–285.

[32] Yule G. 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, 1–64.