Reduced Rank of Regression Using Generalized Method of Moments Estimators

Kleibergen, F.

Publication date:
1996

Reduced Rank Regression using Generalized Method of Moments Estimators

with extensions to structural breaks in reduced rank models

Frank Kleibergen*

April 24, 1996

Abstract

Generalized Method of Moments (GMM) Estimators are derived for Reduced Rank Regression Models, the Error Correction Cointegration Model (ECCM) and the Incomplete Simultaneous Equations Model (INSEM). The GMM (2SLS) estimators of the cointegrating vector in the ECCM are shown to have normal limiting distributions. Tests for the number of unit roots can be constructed straightforwardly and have Dickey-Fuller type limiting distributions. Two extensions of the ECCM, which are important in practice, are analyzed. First, cointegration estimators and tests allowing for structural shifts in the variance (heteroscedasticity) of the series are derived and analyzed using both a Generalized Least Squares Estimator and a White Covariance Matrix Estimator. The resulting cointegrating vectors estimators have again normal limiting distributions while the cointegration tests have identical limiting distributions which differ from the Dickey-Fuller type. Second, cointegrating vector estimators and tests are derived which allow for structural breaks in the cointegrating vector and/or multiplicator. The limiting distributions of the estimators are again shown to be normal and the limiting distributions of the cointegration tests differ from the Dickey-Fuller type.

*Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, Email:f.r.kleibergen@kub.nl

1 Introduction

Cointegration has been an important research topic since its definition in [1] and already a large literature has evolved on it. An important part of this literature is devoted to the construction of estimators, test statistics and their limiting distributions, see a.o. [1], [4] and [11]. These contributions cover stylized models, which are constant over time and have a constant variance. Although models, which deviate from these assumptions, no longer suffice the condition of weak (covariance) stationarity, they can still show mean reversion so that they still possess properties of cointegration, see for example [7], where it is shown that cointegration can still be defined in periodic models although the model for the cointegrating relationships is not weakly stationary but still mean reverting. So, the cointegrating relationships do not suffice weak stationarity conditions in these cases but cointegration is still an important property of the series generated by these kind of models. In practice, there is a need for the construction of cointegration estimators and test statistics, which can be applied in these kind of models as a large number of series possess properties resulting from these models, like heteroscedasticity and structural breaks, and still show mean reversion of linear combinations. Examples are numerous and lie especially in areas like finance, where heteroscedasticity is a stylized fact, and macro-economics, where structural breaks are an important topic. Naive application of the cointegration estimators, which essentially assume that these properties are not present, can lead to inconsistent estimators and/or wrong expressions of the (asymptotic) variances of the estimators. There is, therefore, a need for the development of estimators and test statistics, which can be applied in these kind of models. This paper tries to contribute to this topic by developing a Generalized Method of Moments (GMM) framework, see [3], for cointegration models, which allows for the incorportation of for example heteroscedasticity and/or structural breaks. Also the stylized models are covered by this framework and lead to estimators, which are the 2SLS (two stage least squares) counterpart of the canonical correlation cointegration estimators, see [4].

The discussion of this GMM framework for cointegration analysis, is organized as follows. In section 2, the relation between the 2SLS estimators in cointegration and simultaneous equations models is discussed jointly with the limiting distributions of the cointegrating vector estimators for a few widely used specifications of the deterministic components. Section 3, contains a discussion of a GMM statistic (=GMM objective function) for testing for the number of unit roots/cointegrating relationships. In section 4, the stylized model is extended to a model where a shift of variance occurs after a predefined fraction of time has evolved. Both a Generalized Least Squares

approach, which assumes a priori knowledge of the variance shift moment, and a nonparametric approach using a White covariance matrix, see [15], which uses no knowledge about the specification of heteroscedasticity, for the construction of cointegration estimators and statistics that account for heteroscedasticity are discussed. In section 5, cointegration estimators and statistics that account for a change in the cointegrating relationship and/or multiplicator, are constructed. Both extensions can be further generalized to more shifts and also other moment conditions can be added. Finally, the sixth section concludes.

Note that the following definitions are used throughout the paper; $\Rightarrow$ indicates weak convergence; integrals are taken over the unit interval unless indicated otherwise; when possible without confusion, integrals like $\int W(t)dt$ are shortly denoted as $\int W$. The theorems in the paper are derived assuming Gaussian disturbances, which assumption can be relaxed considerably, see for example [14].

2 2SLS Estimators in reduced rank regression models

2.1 Reduced Rank Regression Models

Reduced rank regression models are characterized by the lower column or row rank of a parameter matrix. Two well known models which possess this property are the Error Correction Cointegration Model (ECCM) and the INcomplete Simultaneous Equations Model (INSEM). The ECCM is specified as

$$\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t, \tag{1}$$

where $x_t : k \times 1$, $t = 1, ..., T$; $\alpha, \beta : k \times r$; $\beta' = (I_r \text{ -}\beta_2')$; and ε_t is Gaussian white noise with covariance matrix Σ. For simplicity higher order lags are left out. The INSEM reads

$$\begin{aligned} y_{1t} &= \beta_2' y_{2t} + \gamma_1' x_{1t} + \varepsilon_{1t} \\ y_{2t} &= \Pi_{21} x_{1t} + \Pi_{22} x_{2t} + \varepsilon_{2t} \end{aligned} \tag{2}$$

where $y_{1t} : m_1 \times 1$, $y_{2t} : m_2 \times 1$, $x_{1t} : k_1 \times 1$, $x_{2t} : k_2 \times 1$, $t = 1, ..., T$; $\beta_2 : m_2 \times m_1$; $\gamma_1 : k_1 \times m_1$; $\Pi_{21} : m_2 \times k_1$; $\Pi_{22} : m_2 \times k_2$. The disturbances ε_{1t} and ε_{2t} are assumed to be Gaussian white noise with covariance matrix Σ. The variables x_{1t} and x_{2t} are assumed to be (weakly) exogenous. The INSEM in equation (2) is identified when the number of excluded exogenous variables from the first set of equations, k_2, is at least as large as the number

of equations in the second set, m_2, $k_2 \geq m_2$. The reduced rank property of both models is obtained when we specify a general model,

$$z_t = \Pi w_t + u_t. \tag{3}$$

Both the ECCM and the INSEM are restricted versions of the model in equation (3). The ECCM is obtained by specifying $z_t = \Delta x_t$, $\Pi = \alpha\beta' = \begin{pmatrix} \alpha_{11} & -\alpha_{11}\beta_2' \\ \alpha_{21} & -\alpha_{21}\beta_2' \end{pmatrix}$, $u_t = \varepsilon_t$, $w_t = x_{t-1}$, while the INSEM is obtained when $z_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$, $w_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix}$, $u_t = \begin{pmatrix} \varepsilon_{1t} + \beta_2'\varepsilon_{2t} \\ \varepsilon_{2t} \end{pmatrix}$, $\Pi = \begin{pmatrix} \beta_2'\Pi_{21} + \gamma_1 & \beta_2'\Pi_{22} \\ \Pi_{21} & \Pi_{22} \end{pmatrix}$. The reduced rank structure of the ECCM is obvious while the INSEM has a reduced rank structure when $\gamma_1 = 0$ since the first set of rows of Π is a linear function of the other rows in that case. The reduced rank properties of both models are different in nature, however, as in the ECCM the last set of columns is a linear combination of the first set while in the INSEM the first set of rows is a linear combination of the last set.

2.2 2SLS estimators

In the INSEM from equation (2) a consistent 2SLS estimator of the parameters β_2 and γ_1 is obtained when we replace Π_{21} and Π_{22} by their least squares estimators obtained from the second set of equations. A similar kind of 2SLS estimator can be constructed for the cointegrating vector β in the ECCM. An important difference between the cointegrating vector parameter β and the structural form parameters β_2 and γ_1 results, however, from the presence of the cointegrating vector in all equations of the ECCM while the structural form parameters of the INSEM only appear in the first set of equations. The 2SLS estimator for the ECCM has, therefore, a completely different specification then the 2SLS estimator in the INSEM. Both estimators can be derived in a Generalized Method of Moments (GMM) framework, see [3].

To derive the expressions of the 2SLS estimators both in the INSEM as the ECCM, we use the first order conditions for a maximum of the likelihood. The derivatives of the log likelihood, when assuming Gaussian white noise disturbances with covariance matrix Σ, of the model in equation (3), read

$$\begin{aligned} \frac{\partial \ln l(\theta)}{\partial \theta'} &= vec(\Sigma^{-1})' \sum_{t=1}^{T} (u_t \otimes I_k) \frac{\partial u_t}{\partial \theta'} \\ &= \sum_{t=1}^{T} vec(\Sigma^{-1} u_t)' \frac{\partial u_t}{\partial \theta'} \end{aligned} \tag{4}$$

$$
\begin{aligned}
&= \sum_{t=1}^{T} vec(\Sigma^{-1}u_t)'(w_t' \otimes I_k)\frac{\partial vec(\Pi)}{\partial \theta'} \\
&= \sum_{t=1}^{T} vec(u_t w_t')'(I_k \otimes \Sigma^{-1})\frac{\partial vec(\Pi)}{\partial \theta'}
\end{aligned}
$$

In the GMM objective function we will only use the $\sum_{t=1}^{T} vec(u_t w_t')$ part of the derivative in equation (4). When we substitute the correct expression for θ in $\frac{\partial vec(\Pi)}{\partial \theta'}$, the first order derivatives of the different parameters are obtained. These expressions read, for the ECCM,

$$
\begin{aligned}
\frac{\partial vec(\Pi)}{\partial vec(\beta')'} &= -(I_k \otimes \alpha), \qquad (5)\\
\frac{\partial vec(\Pi)}{\partial vec(\alpha)'} &= -(\beta \otimes I_k),
\end{aligned}
$$

and for the INSEM,

$$
\begin{aligned}
\frac{\partial vec(\Pi)}{\partial vec(\beta_2')'} &= -(\begin{pmatrix} \Pi_{21} & \Pi_{22} \end{pmatrix}' \otimes \begin{pmatrix} I_{m_1} \\ 0 \end{pmatrix}), \qquad (6)\\
\frac{\partial vec(\Pi)}{\partial vec(\gamma_1')'} &= -(\begin{pmatrix} I_{k_1} & 0 \end{pmatrix}' \otimes \begin{pmatrix} I_{m_1} \\ 0 \end{pmatrix}),\\
\frac{\partial vec(\Pi)}{\partial vec(\Pi_{21})'} &= -(\begin{pmatrix} I_{k_1} & 0 \end{pmatrix}' \otimes \begin{pmatrix} \beta_2' \\ I_{m_2} \end{pmatrix}),\\
\frac{\partial vec(\Pi)}{\partial vec(\Pi_{22})'} &= -(\begin{pmatrix} 0 & I_{k_2} \end{pmatrix}' \otimes \begin{pmatrix} \beta_2' \\ I_{m_2} \end{pmatrix}).
\end{aligned}
$$

The expressions of the derivatives of the individual parameters are substituted in the first order derivative of the objection function which is minimized in the GMM framework. As we cannot exactly solve for the normal equation, $\sum_{t=1}^{T} u_t w_t' = 0$, in case of reduced rank parameter matrices, we take a quadratic form containing these normal equations as objective function to be minimized in the GMM framework, see also [2],

$$
G(\theta) = vec(\sum_{t=1}^{T} u_t w_t')'((\sum_{t=1}^{T} w_t w_t')^{-1} \otimes \Sigma^{-1})vec(\sum_{t=1}^{T} u_t w_t'). \qquad (7)
$$

The first order conditions of the GMM objective function then become

$$
\frac{\partial G(\theta)}{\partial \theta'} = 0 \Leftrightarrow \quad (8)
$$

$$\sum_{t=1}^{T}(\frac{\partial u_t}{\partial \theta'})(w_t' \otimes I_k)((\sum_{t=1}^{T} w_t w_t')^{-1} \otimes \Sigma^{-1})vec(\sum_{t=1}^{T} u_t w_t') = 0 \Leftrightarrow$$

$$(\frac{\partial vec(\Pi)}{\partial \theta'})'(\sum_{t=1}^{T} w_t w_t' \otimes I_k)((\sum_{t=1}^{T} w_t w_t')^{-1} \otimes \Sigma^{-1})vec(\sum_{t=1}^{T} u_t w_t') = 0 \Leftrightarrow$$

$$(\frac{\partial vec(\Pi)}{\partial \theta'})'(I_k \otimes \Sigma^{-1})vec(\sum_{t=1}^{T} u_t w_t') = 0 \Leftrightarrow$$

$$(\frac{\partial vec(\Pi)}{\partial \theta'})'vec(\Sigma^{-1}\sum_{t=1}^{T} u_t w_t') = 0$$

The first order condition of the GMM objective function in equation (8) exactly equals the first order condition for a maximum likelihood value, see equation (4).

For the different parameters of the ECCM these first order conditions read,

$$(I_k \otimes \alpha)'vec(\Sigma^{-1}\sum_{t=1}^{T} u_t x_{t-1}') = 0 \Leftrightarrow \tag{9}$$

$$vec(\alpha'\Sigma^{-1}\sum_{t=1}^{T}(\Delta x_t - \alpha\beta' x_{t-1})x_{t-1}') = 0 \Leftrightarrow$$

$$(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1}(\sum_{t=1}^{T} x_{t-1}\Delta x_t')\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} = \beta,$$

$$(\beta \otimes I_k)'vec(\Sigma^{-1}\sum_{t=1}^{T} u_t x_{t-1}') = 0 \Leftrightarrow \tag{10}$$

$$vec(\Sigma^{-1}\sum_{t=1}^{T}(\Delta x_t - \alpha\beta' x_{t-1})x_{t-1}'\beta) = 0 \Leftrightarrow$$

$$(\sum_{t=1}^{T}\Delta x_t x_{t-1}'\beta)(\beta'\sum_{t=1}^{T} x_{t-1}x_{t-1}'\beta)^{-1} = \alpha.$$

For the different parameters of the INSEM these first order conditions read,

$$(\begin{pmatrix} \Pi_{21} & \Pi_{22} \end{pmatrix}' \otimes \begin{pmatrix} I_{m_1} \\ 0 \end{pmatrix})'vec(\Sigma^{-1}\sum_{t=1}^{T} u_t x_t') = 0 \Leftrightarrow \tag{11}$$

$$vec(\sum_{t=1}^{T}(y_{1t} - \beta_2' y_{2t} - \gamma_1' x_{1t})x_t'\begin{pmatrix} \Pi_{21} & \Pi_{22} \end{pmatrix}') = 0 \Leftrightarrow$$

$$(\begin{pmatrix} \Pi_{21} & \Pi_{22} \end{pmatrix}\sum_{t=1}^{T} x_t y_{2t}')^{-1}\begin{pmatrix} \Pi_{21} & \Pi_{22} \end{pmatrix}\sum_{t=1}^{T} x_t(y_{1t} - \gamma_1' x_{1t})' = \beta_2,$$

$$
\begin{aligned}
(\begin{pmatrix} I_{k_1} & 0 \end{pmatrix}' \otimes \begin{pmatrix} I_{m_1} \\ 0 \end{pmatrix})' vec(\Sigma^{-1} \sum_{t=1}^{T} u_t x_t') &= 0 \Leftrightarrow \qquad (12) \\
vec(\sum_{t=1}^{T} (y_{1t} - \beta_2' y_{2t} - \gamma_1' x_{1t}) x_{1t}') &= 0 \Leftrightarrow \\
(\sum_{t=1}^{T} x_{1t} x_{1t}')^{-1} \sum_{t=1}^{T} x_{1t} (y_{1t} - \beta_2' y_{2t})' &= \gamma_1,
\end{aligned}
$$

$$
\begin{aligned}
(I_k \otimes \begin{pmatrix} \beta_2' \\ I_{m_2} \end{pmatrix})' vec(\Sigma^{-1} \sum_{t=1}^{T} u_t x_t') &= 0 \Leftrightarrow \qquad (13) \\
vec(\begin{pmatrix} \beta_2' \\ I_{m_2} \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} I_{m_1} & \beta_2' \\ 0 & I_{m_2} \end{pmatrix} \sum_{t=1}^{T} \varepsilon_t x_t') &= 0 \Leftrightarrow \\
vec(\begin{pmatrix} 0 \\ I_{m_2} \end{pmatrix}' \Omega^{-1} \sum_{t=1}^{T} \varepsilon_t x_t') &= 0 \Leftrightarrow \\
vec(\sum_{t=1}^{T} (y_{2t} - \Pi_2 x_t) x_t') &= 0 \Leftrightarrow \\
(\sum_{t=1}^{T} y_{2t} x_t')(\sum_{t=1}^{T} x_t x_t')^{-1} &= \Pi_2
\end{aligned}
$$

where $\Omega = \begin{pmatrix} I_{m_1} & \beta_2' \\ 0 & I_{m_2} \end{pmatrix}' \Sigma \begin{pmatrix} I_{m_1} & \beta_2' \\ 0 & I_{m_2} \end{pmatrix}$. The normal equations for the INSEM directly lead to the well known 2SLS estimator for INSEMs as the estimator of Π_2 is independent of the parameters β_2 and γ_1 such that it can be estimated independently. The resulting estimate of Π_2 is then used to construct estimators for β_2 and γ_1 (2SLS estimators). The estimators of α and β in the ECCM both depend on one another. As we didnot restrict α and β, they are also not identified. If we specify β as, $\beta = \begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix}$, both α and β_2 are properly identified. If this specification of β is used, a consistent estimator of α is,

$$
\begin{aligned}
\alpha \;=\; & (\sum_{t=1}^{T} \Delta x_t (1 - x_{2t-1}' (\sum_{t=1}^{T} x_{2t-1} x_{2t-1}')^{-1} x_{2t-1}) x_{1t-1}') \qquad (14) \\
& (\sum_{t=1}^{T} x_{1t-1} (1 - x_{2t-1}' (\sum_{t=1}^{T} x_{2t-1} x_{2t-1}')^{-1} x_{2t-1}) x_{1t-1}')^{-1}
\end{aligned}
$$

where $x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix}$, $x_{1t} : r \times 1$, $x_{2t} : (k-r) \times 1$. If the estimator of α from equation (14) is used in the estimation of the cointegrating vector parameter

β, equation (9), the identifying restrictions on β are automatically fulfilled. The resulting estimator of β is then the 2SLS estimator of the cointegrating vector β. In a Bayesian analysis this 2SLS estimator equals to the mean of the conditional posterior of β given α when a diffuse prior is used, see [6]. The estimators of α and β in equations (9) and (10) also allow for the construction of an iterative estimation scheme for which the resulting estimators converge to the maximum likelihood estimators. Asymptotically the 2SLS least squares cointegrating vector estimator possesses the same kind of properties as the maximum likelihood estimator, i.e. superconsistency and normal limiting distribution. This is proved in the theorems in the following section.

2.3 Limiting distributions 2SLS estimators

As the limiting distribution of the 2SLS estimator in the INSEM model is discussed at length in the literature, see for example [9], we concentrate on the limiting distribution of the 2SLS estimator for the cointegration case. Theorem 1 states the limiting distribution of the multiplicator estimator, $\hat{\alpha}$, and the 2SLS cointegrating vector estimator, $\hat{\beta}$.

Theorem 1 *When the DataGenerating Process (DGP) in equation (1) is such that the number of cointegrating vectors equals r ($k-r$ unit roots), the estimators*

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - x'_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1}) \tag{15}$$
$$(\sum_{t=1}^{T} x_{1t-1}(1 - x'_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1})^{-1}$$

and

$$\hat{\beta} = (\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{T} x_{t-1}\Delta x'_t)\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \tag{16}$$

have a limiting behavior which can be characterized by

$$\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow n(0,\ cov(\beta' x)^{-1} \otimes \Sigma)) \tag{17}$$

$$T(\hat{\beta} - \beta) \Rightarrow \begin{pmatrix} 0 \\ (\beta'_\perp \beta_\perp)^{-1}\beta'_\perp \alpha_\perp \Lambda_1^{-1\prime}(\int W_1 W'_1)^{-1} \int W_1 dW'_2 \Lambda'_2 \end{pmatrix} \tag{18}$$
$$\Rightarrow \begin{pmatrix} 0 \\ n(0, \alpha'\Sigma^{-1}\alpha \otimes \Theta \end{pmatrix},$$

where W_1, resp. W_2 are $(k-r)$, resp. r dimensional stochastically independent Brownian motions defined on the unit interval, $\Lambda_1 = (\alpha'_\perp \Sigma \alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2 = (\alpha' \Sigma^{-1} \alpha)^{\frac{1}{2}}$, $\Theta = (\beta'_\perp \beta_\perp)^{-1} \beta'_\perp \alpha_\perp \Lambda_1^{-1\prime} (\int W_1 W_1')^{-1} \Lambda_1^{-1} \alpha'_\perp \beta_\perp (\beta'_\perp \beta_\perp)^{-1}$ and Σ is estimated by the sum of squared residuals, $\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} (\Delta x_t (1 - x'_{t-1} (\sum_{t=1}^{T} x_{t-1} x'_{t-1})^{-1} x_{t-1}) \Delta x'_t)$.

Proof: see appendix.

Theorem 1 discusses the limiting distribution of the cointegrating vector estimator for the most straightforward case, i.e. no further lags in the VAR polynomial and no deterministic components, and shows that it is identical to the limiting distribution of the canonical correlation maximum likelihood estimator, see [4]. While addition of lags of Δx_t only changes the limiting distribution of the cointegrating vector estimator, $\hat{\beta}$, in the sense that $\alpha'_\perp \beta_\perp$ has to be replaced by $\alpha'_\perp \Gamma(1) \beta_\perp$, where $\Gamma(L) \Delta x_t = \alpha \beta' x_{t-1} + \varepsilon_t$, $\Gamma(L)$ is a $(p-1)$-dimensional lag polynomial in case of a VAR(p), inclusion of deterministic components does also change the functional form of the cointegrating vector estimator, see for example [4] and [5] for the influence of the deterministic components on other kind of cointegrating vector estimators. Theorem 2 states the estimators and limiting distributions of the multiplicator and cointegrating vector estimators including deterministic components for a few widely used specifications of the deterministic components.

Theorem 2 *When the DGP reads*

$$\Delta x_t = \alpha(\beta' x_{t-1} + \mu') + \varepsilon_t, \tag{19}$$

and the number of cointegrating vectors equals r ($k-r$ unit roots), the estimators

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1} \tag{20}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}) x'_{1t-1}) (\sum_{t=1}^{T} x_{1t-1} (1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}) x'_{1t-1})^{-1}$$

and

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \end{pmatrix} = (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1} (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \Delta x'_t) \Sigma^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma^{-1} \hat{\alpha})^{-1} \tag{21}$$

have a limiting behavior which can be characterized by

$$\sqrt{T}(\hat{\alpha}-\alpha) \Rightarrow n(0,\ cov(\beta' x-\mu')^{-1}\otimes\Sigma) \tag{22}$$

$$\begin{pmatrix} TI_k & 0 \\ 0 & T^{\frac{1}{2}} \end{pmatrix}\begin{pmatrix} \hat{\beta}-\beta \\ \hat{\mu}-\mu \end{pmatrix} \tag{23}$$

$$\Rightarrow \begin{pmatrix} \begin{pmatrix} (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\Lambda_1^{-1\prime} & 0 \\ 0 & 1 \end{pmatrix}(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix}\begin{pmatrix} W_1 \\ \iota \end{pmatrix}')^{-1}(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix} dW_2')\Lambda_2' \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 0 \\ n(0,\alpha'\Sigma^{-1}\alpha\otimes\Theta_1 \end{pmatrix},$$

When the DGP reads

$$\Delta x_t = c+\alpha\beta' x_{t-1}+\varepsilon_t, \tag{24}$$

$c=\alpha\mu'+\alpha_\perp\lambda'$, *and the number of cointegrating vectors equals* r $(k-r$ *unit roots), the estimators*

$$\hat{\alpha} = (\sum_{t=1}^{T}\Delta x_t(1-\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1} \tag{25}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}')(\sum_{t=1}^{T}x_{1t-1}(1-\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}')^{-1}$$

and

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \end{pmatrix} = (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1}(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\Delta x_t')\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \tag{26}$$

have a limiting behavior which can be characterized by

$$\sqrt{T}(\hat{\alpha}-\alpha) \Rightarrow n(0,\ cov(\beta' x-\mu')^{-1}\otimes\Sigma)) \tag{27}$$

$$\begin{pmatrix} \begin{pmatrix} TI_{k-r-1} & 0 \\ 0 & T^{\frac{3}{2}} \end{pmatrix}(\beta_{2\perp}^{*\prime}\beta_{2\perp}^{*})^{-1}\beta_{2\perp}^{*\prime} & 0 \\ 0 & T^{\frac{1}{2}} \end{pmatrix}\begin{pmatrix} \beta_2-\hat{\beta}_2 \\ \hat{\mu}-\mu \end{pmatrix} \tag{28}$$

$$\Rightarrow \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1\prime}(\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}\begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}')^{-1}\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} dW_2'\Lambda_2'$$

$$\Rightarrow n(0,\ \alpha'\Sigma^{-1}\alpha\otimes\Theta_2)$$

When the DGP reads

$$\Delta x_t = c + \alpha(\beta' x_{t-1} + \delta' t) + \varepsilon_t, \tag{29}$$

$c = \alpha\mu' + \alpha_\perp\lambda'$, *and the number of cointegrating vectors equals* r *(*$k-r$ *unit roots), the estimators*

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}')^{-1} \tag{30}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}) x'_{1t-1}) (\sum_{t=1}^{T} x_{1t-1} (1 - \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}')^{-1}$$
$$\begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}) x'_{1t-1})^{-1}$$

and

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \\ \hat{\delta} \end{pmatrix} = (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \\ t \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \\ t \end{pmatrix}')^{-1} (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \\ t \end{pmatrix} \Delta x'_t) \Sigma^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma^{-1} \hat{\alpha})^{-1} \tag{31}$$

have a limiting behavior which can be characterized by

$$\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow n(0,\ cov(\beta' x - \mu' - \delta' t)^{-1} \otimes \Sigma)) \tag{32}$$

$$\begin{pmatrix} TI_k & 0 & 0 \\ 0 & T^{\frac{1}{2}} & 0 \\ 0 & 0 & T^{\frac{3}{2}} \end{pmatrix} \begin{pmatrix} \hat{\beta} - \beta \\ \hat{\mu} - \mu \\ \hat{\delta} - \delta \end{pmatrix} \tag{33}$$

$$\Rightarrow \begin{pmatrix} & & 0 \\ \begin{pmatrix} (\beta'_\perp \beta_\perp)^{-1} \beta'_\perp \alpha_\perp \Lambda_1^{-1\prime} & 0 \\ 0 & I_2 \end{pmatrix} & & (\int \begin{pmatrix} W_1 \\ \iota \\ \tau \end{pmatrix} \begin{pmatrix} W_1 \\ \iota \\ \tau \end{pmatrix}')^{-1} (\int \begin{pmatrix} W_1 \\ \iota \\ \tau \end{pmatrix}' dW_2') \Lambda_2' \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 0 \\ n(0, \alpha' \Sigma^{-1} \alpha \otimes \Theta_3 \end{pmatrix},$$

where W_1, W_{11} *and* W_2 *are* $(k-r)$, $(k-r-1)$ *and* r *dimensional stochastically independent Brownian motions,* $\Lambda_1 = (\alpha'_\perp \Sigma \alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2 = (\alpha' \Sigma^{-1} \alpha)^{\frac{1}{2}}$, $\Lambda_3 = (\begin{pmatrix} \lambda_\perp (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \Sigma^{\frac{1}{2}} \\ \lambda\lambda' \end{pmatrix} \begin{pmatrix} \lambda_\perp (\alpha'_\perp \alpha_\perp)^{-1} \alpha'_\perp \Sigma^{\frac{1}{2}} \\ \lambda\lambda' \end{pmatrix}')^{\frac{1}{2}}$, $\tau(t) = t$, $\iota(t) = 1$, $0 \leq t \leq 1$, $\beta = \begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix}$, $\beta^*_{2\perp} = (\beta'_\perp \beta_\perp)^{-1} \beta'_\perp \alpha_\perp (\alpha'_\perp \alpha_\perp)^{-1} \begin{pmatrix} \lambda_\perp \\ \lambda \end{pmatrix}'$,

$$\Theta_1 = (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\Lambda_1^{-1\prime}(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix}\begin{pmatrix} W_1 \\ \iota \end{pmatrix}')^{-1}\Lambda_1^{-1}\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1},$$

$$\Theta_2 = \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1\prime}(\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}\begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}')^{-1}\begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1},$$

$$\Theta_3 = (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\Lambda_1^{-1\prime}(\int \begin{pmatrix} W_1 \\ \iota \\ \tau \end{pmatrix}\begin{pmatrix} W_1 \\ \iota \\ \tau \end{pmatrix}')^{-1}\Lambda_1^{-1}\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}.$$

Proof: the first and third part of the theorem are natural extensions of theorem 1. The second part of the theorem is proved in the appendix.

Theorems 1 and 2 show that the limiting distributions of elements of the cointegrating vector estimator are normal and standard (asymptotic) χ^2 tests can be performed to test hypotheses on the cointegrating vectors, see [11]. The next section discusses the use of the cointegrating vector estimator, $\hat{\beta}$, and the multiplicator, $\hat{\alpha}$, in the GMM objective function, equation (7), to construct a statistic to test for the number of cointegrating vectors, unit roots, in the system.

3 Cointegration testing using 2SLS estimators

The GMM objective function, equation (7), can also be used to test for the number of cointegrating vectors, unit roots. This can be done as the optimal value of the objective function has a specific kind of distribution under $H_0 : r = r^*$. In theorem 3, the functional expressions of this objective function for several specifications of the deterministic components and their limiting distributions are stated.

Theorem 3 *When the DGP reads,*

$$\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t, \tag{34}$$

and the number of cointegrating vectors equals r (k-r) unit roots, the use of the estimators,

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - x'_{2t-1} (\sum_{t=1}^{T} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1}) \tag{35}$$
$$(\sum_{t=1}^{T} x_{1t-1} (1 - x'_{2t-1} (\sum_{t=1}^{T} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1})^{-1},$$

and

$$\hat{\beta} = (\sum_{t=1}^{T} x_{t-1} x'_{t-1})^{-1} (\sum_{t=1}^{T} x_{t-1} \Delta x'_t) \Sigma^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma^{-1} \hat{\alpha})^{-1}, \tag{36}$$

in the optimal value of the GMM objective function, equation (7),

$$G(\hat{\alpha}, \hat{\beta}) = vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha} \hat{\beta}' x_{t-1}) x'_{t-1})' \tag{37}$$
$$((\sum_{t=1}^{T} x_{t-1} x'_{t-1})^{-1} \otimes \Sigma^{-1}) vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha} \hat{\beta}' x_{t-1}) x'_{t-1}),$$

leads to a limiting behavior of this optimal value which can be characterized by

$$G(\hat{\alpha}, \hat{\beta}) \tag{38}$$
$$\Rightarrow tr[(\int W_1 dW'_1)' (\int W_1 W'_1)^{-1} (\int W_1 dW'_1)].$$

When the DGP reads,

$$\Delta x_t = \alpha(\beta' x_{t-1} - \mu') + \varepsilon_t, \tag{39}$$

and the number of cointegrating vectors equals r (k-r) unit roots, the use of the estimators,

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}) x'_{1t-1}) \tag{40}$$
$$(\sum_{t=1}^{T} x_{1t-1} (1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1} \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}) x'_{1t-1})^{-1},$$

and

$$\begin{pmatrix} \hat{\beta} \\ -\hat{\mu} \end{pmatrix} = (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1} (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \Delta x'_t) \Sigma^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma^{-1} \hat{\alpha})^{-1}, \tag{41}$$

in the optimal value of the GMM objective function, equation (7),

$$G(\hat{\alpha},\hat{\beta},\hat{\mu}) = vec(\sum_{t=1}^{T}(\Delta x_t - \hat{\alpha}(\hat{\beta}' x_{t-1} - \hat{\mu}'))\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')' \quad (42)$$
$$((\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma^{-1})vec(\sum_{t=1}^{T}(\Delta x_t - \hat{\alpha}(\hat{\beta}' x_{t-1} - \hat{\mu}'))\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}'),$$

leads to a limiting behavior of this optimal value which can be characterized by

$$G(\hat{\alpha},\hat{\beta},\hat{\mu}) \quad (43)$$
$$\Rightarrow tr[(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix} dW_1')'(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix}\begin{pmatrix} W_1 \\ \iota \end{pmatrix}')^{-1}(\int \begin{pmatrix} W_1 \\ \iota \end{pmatrix} dW_1')].$$

When the DGP reads,

$$\Delta x_t = c + \alpha\beta' x_{t-1} + \varepsilon_t, \quad (44)$$

and the number of cointegrating vectors equals r (k-r) unit roots, the use of the estimators,

$$\hat{\alpha} = (\sum_{t=1}^{T}\Delta x_t(1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}'(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}')(45)$$
$$(\sum_{t=1}^{T} x_{1t-1}(1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}'(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}')^{-1},$$

$$\hat{c} = (\sum_{t=1}^{T}\Delta x_t(1 - x_{t-1}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1}x_{t-1})) \quad (46)$$
$$(\sum_{t=1}^{T}(1 - x_{t-1}{}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}{}')^{-1}x_{t-1})^{-1},$$

and

$$\hat{\beta} = (\sum_{t=1}^{T}\bar{x}_{t-1}\bar{x}_{t-1}')^{-1}(\sum_{t=1}^{T}\bar{x}_{t-1}\Delta x_t')\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}, \quad (47)$$

in the optimal value of the GMM objective function, equation (7),

$$G(\hat{\alpha},\hat{\beta},\hat{c}) = vec(\sum_{t=1}^{T}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1} - \hat{c}))\bar{x}_{t-1}')' \quad (48)$$
$$((\sum_{t=1}^{T}\bar{x}_{t-1}\bar{x}_{t-1}')^{-1} \otimes \Sigma^{-1})vec(\sum_{t=1}^{T}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1} - \hat{c}))\bar{x}_{t-1}'),$$

leads to a limiting behavior of this optimal value which can be characterized by

$$G(\hat{\alpha},\hat{\beta},\hat{c}) \quad (49)$$
$$\Rightarrow tr[(\int \begin{pmatrix} \bar{W}_{11} \\ \bar{\tau} \end{pmatrix} dW_1')'(\int \begin{pmatrix} \bar{W}_{11} \\ \bar{\tau} \end{pmatrix} \begin{pmatrix} \bar{W}_{11} \\ \bar{\tau} \end{pmatrix}')^{-1}(\int \begin{pmatrix} \bar{W}_{11} \\ \bar{\tau} \end{pmatrix} dW_1')].$$

When the DGP reads,

$$\Delta x_t = c + \alpha(\beta' x_{t-1} - \delta' t) + \varepsilon_t, \quad (50)$$

and the number of cointegrating vectors equals r (k-r) unit roots, the use of the estimators,

$$\hat{\alpha} = (\sum_{t=1}^{T} \Delta x_t (1 - \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}')^{-1} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}) x_{1t-1}') \quad (51)$$
$$(\sum_{t=1}^{T} x_{1t-1}(1 - \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}')^{-1} \begin{pmatrix} x_{2t-1} \\ 1 \\ t \end{pmatrix}) x_{1t-1}')^{-1},$$

$$\hat{c} = (\sum_{t=1}^{T} \Delta x_t (1 - \begin{pmatrix} x_{t-1} \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix}')^{-1} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix})) \quad (52)$$
$$(\sum_{t=1}^{T} (1 - \begin{pmatrix} x_{t-1} \\ t \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix}')^{-1} \begin{pmatrix} x_{t-1} \\ t \end{pmatrix})^{-1},$$

and

$$\begin{pmatrix} \hat{\beta} \\ -\hat{\delta} \end{pmatrix} = (\sum_{t=1}^{T} \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix} \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix}')^{-1} (\sum_{t=1}^{T} \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix} \Delta x_t') \Sigma^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma^{-1} \hat{\alpha})^{-1}, \quad (53)$$

in the optimal value of the GMM objective function, equation (7),

$$G(\hat{\alpha},\hat{\beta},\hat{\delta},\hat{c}) = vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha}(\hat{\beta}' x_{t-1} - \hat{\delta}' t) - \hat{c}) \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix}')' \quad (54)$$
$$((\sum_{t=1}^{T} \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix} \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix}')^{-1} \otimes \Sigma^{-1})$$
$$vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha}(\hat{\beta}' x_{t-1} - \hat{\delta} t') - \hat{c}) \begin{pmatrix} \bar{x}_{t-1} \\ \bar{t} \end{pmatrix}'),$$

leads to a limiting behavior of this optimal value which can be characterized by

$$G(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{c}) \qquad (55)$$

$$\Rightarrow tr[(\int \begin{pmatrix} \bar{W}_1 \\ \bar{\tau} \end{pmatrix} dW_1')'(\int \begin{pmatrix} \bar{W}_1 \\ \bar{\tau} \end{pmatrix} \begin{pmatrix} \bar{W}_1 \\ \bar{\tau} \end{pmatrix}')^{-1}(\int \begin{pmatrix} \bar{W}_1 \\ \bar{\tau} \end{pmatrix} dW_1')].$$

where $x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix}$, $x_{1t} : r \times 1$, $x_{2t} : (k-r) \times 1$; $\bar{x}_{t-1} = x_{t-1} - \frac{1}{T}\sum_{t=1}^{T} x_{t-1}$, $\bar{t} = t - \frac{1}{T}\sum_{t=1}^{T} t$; W_1, W_{11} *are* $(k-r)$, $(k-r-1)$ *dimensional brownian motions,* $W_1 = \begin{pmatrix} W_{11} \\ w_{12} \end{pmatrix}$, $\bar{W}_1 = W_1 - \int W_1$, $\bar{W}_{11} = W_{11} - \int W_{11}$, $\tau(t) = t$, $\iota(t) = 1$, $0 \le t \le 1$, $\bar{\tau} = \tau - \int \tau$, *and* Σ *is estimated by the residual sum of squares for the unrestricted model.*

Proof: for the first part a proof is given in the appendix, the other parts follow naturally.

Theorems 1 to 3 show that the limiting distributions using the 2SLS (GMM) estimators are identical to the limiting distributions when maximum likelihood estimators are used, see [4]. As maximum likelihood estimators can be constructed in a straightforward way using canonical correlations there is not much gain when 2SLS estimators are used compared to maximum likelihood estimators from a limiting distribution perspective. Possible gains can lie both in the small sample distribution of the 2SLS estimator and in model extensions as maximum likelihood estimators become analytically intractable when more complicated models are used then the one shown in equation (1).

In [12], it is shown that the canonical correlation cointegrating vector estimator has a small sample distribution with Cauchy type tails such that it has no finite moments. When we neglect the dynamic property of the data and assume fixed regressors, results from [9] indicate that the small sample distribution of the cointegrating vector estimator has finite moments up to the degree $(k-r)$. This degree is determined by the $(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}$ expression appearing in the cointegrating vector estimator $\hat{\beta}$. As β is specified such that it always has rank r, rank reduction of $\alpha\beta'$ implies that α has a rank smaller than r. In that case $\hat{\alpha}'\Sigma^{-1}\hat{\alpha}$ would not be invertible leading to the fat tails of the small sample distribution. So, cointegration tests essentially test for the rank of α and can be considered as tests for the local identification of β and

are, therefore, comparable with the concentration parameter in the INSEM, see [9].

The maximum likelihood estimator is appealing as it has a very simple expression in the standard case. The relation between maximum likelihood estimators and canonical correlations is, however, lost when extensions of the model are considered. Furthermore, model extensions often lead to analytically intractable maximum likelihood estimators. The GMM framework used in this paper offers a framework which allows for the analytical construction of cointegrating vector estimators for a general class of models. In the next sections two kind of structural break model extensions are analyzed, i.e. structural breaks in the variance (heteroscedasticity) and cointegrating vector and/or multiplicator, whose cointegrating vector maximum likelihood estimators are not of the canonical correlation type.

4 Cointegration in a Model with Heteroscedasticity

Assuming homoscedastic errors in the model from equation (1), the maximum likelihood estimator of the cointegrating vector in the ECCM can be constructed by means of canonical correlations. This estimator has a normal limiting distribution under conditions which are more general than strict homoscedasticity, see [14], where it is for example proved that the weak convergence properties are retained in case of heteroscedasticity with a constant mean of the conditional variance. These weak convergence properties are, however, lost when the mean of the conditional variance changes from period to period. Furthermore, also the relation between the maximum likelihood estimator and canonical correlations is lost in that case. A (3SLS) GMM cointegrating vector estimator can still be constructed when the functional form of the heteroscedasticity is known. It is also possible to perform a quasi-GMM analysis using a White covariance matrix estimator, see [15]. We construct estimators and limiting distributions for both cases for an example of a change of the variance after a predefined period of time T_1 has evolved, such that the analyzed model reads,

$$\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t, \tag{56}$$

where

$$\begin{aligned} cov(\varepsilon_t) &= \Sigma_1, \qquad t = 1, ..., T_1 \\ &= \Sigma_2. \qquad t = T_1 + 1, ..., T \end{aligned} \tag{57}$$

In the next two subsections, the (quasi) GMM cointegration estimators and tests and their limiting distributions are derived using both a Generalized Least Squares (GLS) framework to account for the heteroscedasticity and a White Covariance Matrix Estimator.

4.1 Generalized Least Squares Cointegration Estimators

Assuming that we know the form of heteroscedasticity, a different GMM objective function then equation (7) is used in the construction of the GMM estimators,

$$\begin{aligned} G(\alpha,\beta) &= vec(\sum_{t=1}^{T_1}\Sigma_1^{-1}\varepsilon_t x_{t-1}' + \sum_{t=T_1+1}^{T}\Sigma_2^{-1}\varepsilon_t x_{t-1}')' \\ &\quad (\sum_{t=1}^{T_1}(x_{t-1}x_{t-1}'\otimes\Sigma_1^{-1}) + \sum_{t=T_1+1}^{T}(x_{t-1}x_{t-1}'\otimes\Sigma_2^{-1}))^{-1} \\ &\quad vec(\sum_{t=1}^{T_1}\Sigma_1^{-1}\varepsilon_t x_{t-1}' + \sum_{t=T_1+1}^{T}\Sigma_2^{-1}\varepsilon_t x_{t-1}'). \end{aligned} \tag{58}$$

In the next theorem the GMM estimators and their limiting distributions jointly with the limiting distribution of the optimal value of the GMM objective function are stated.

Theorem 4 *When the DGP in equations (56), (57) is such that the number of cointegrating vectors is r ($k-r$ unit roots), the estimators,*

$$\begin{aligned} vec(\hat{\alpha}) &= ((\sum_{t=1}^{T_1}x_{t-1}x_{t-1}'\otimes\hat{\Sigma}_1^{-1}) + (\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}'\otimes\hat{\Sigma}_2^{-1}))_1^{-1} \\ &\quad vec(\sum_{t=1}^{T_1}\hat{\Sigma}_1^{-1}\Delta x_t x_{t-1}' + \sum_{t=T_1+1}^{T}\hat{\Sigma}_2^{-1}\Delta x_t x_{t-1}'), \end{aligned} \tag{59}$$

and

$$\begin{aligned} vec(\hat{\beta}') &= ((\sum_{t=1}^{T_1}x_{t-1}x_{t-1}'\otimes\hat{\alpha}'\hat{\Sigma}_1^{-1}\hat{\alpha}) + (\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}'\otimes\hat{\alpha}'\hat{\Sigma}_2^{-1}\hat{\alpha}))^{-1} \\ &\quad vec(\sum_{t=1}^{T_1}\hat{\alpha}'\hat{\Sigma}_1^{-1}\Delta x_t x_{t-1}' + \sum_{t=T_1+1}^{T}\hat{\alpha}'\hat{\Sigma}_2^{-1}\Delta x_t x_{t-1}'), \end{aligned} \tag{60}$$

have a limiting behavior which can be characterized by

$$\begin{aligned}\sqrt{T}vec(\hat{\alpha}-\alpha) \quad &\Rightarrow \quad n(0,(w(cov(\beta'x)_1\otimes\alpha'\Sigma_1^{-1}\alpha) \\ &\quad +(1-w)(cov(\beta'x)_2\otimes\alpha'\Sigma_2^{-1}\alpha))^{-1}),\end{aligned} \tag{61}$$

and

$$\begin{aligned}&T[vec(\beta_2-\hat{\beta}_2)] \\ \Rightarrow \quad &((\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\otimes I_r)((\Lambda_1\int_0^w W_1W_1'\Lambda_1'\otimes\alpha'\Sigma_1^{-1}\alpha)+ \\ &(\int_w^1(\Lambda_2W_1(t)+\Lambda_1W_1(w))(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt\otimes\alpha'\Sigma_2^{-1}\alpha))^{-1} \\ &vec[\Omega_1(\int_0^w dW_2W_1')\Lambda_1'+\Omega_2(\int_w^1 dW_2(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)],\end{aligned} \tag{62}$$

The limiting behavior of the optimal value of the GMM objective function, can be characterized by

$$\begin{aligned}&G(\hat{\alpha},\hat{\beta}) \\ \Rightarrow \quad &vec[\Lambda_1(\int_0^w dW_1W_1')\Lambda_1'+\Lambda_2(\int_w^1 dW_1(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)]' \\ &((\Lambda_1\int_0^w W_1W_1'\Lambda_1'\otimes\alpha_\perp'\Sigma_1^{-1}\alpha_\perp)+(\int_w^1(\Lambda_2W_1(t)+\Lambda_1W_1(w)) \\ &(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt\otimes\alpha_\perp'\Sigma_2^{-1}\alpha_\perp))^{-1} \\ &vec[\Lambda_1(\int_0^w dW_1W_1')\Lambda_1'+\Lambda_2(\int_w^1 dW_1(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)],\end{aligned} \tag{63}$$

where $w=\frac{T_1}{T}$, W_1 *and* W_2, *are stochastically independent* r, $(k-r)$ *dimensional Brownian motions with identity covariance matrices,* $\Lambda_1=(\alpha_\perp'\Sigma_1^{-1}\alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2=(\alpha_\perp'\Sigma_2^{-1}\alpha_\perp)^{\frac{1}{2}}$, $\Omega_1=(\alpha'\Sigma_1\alpha)^{\frac{1}{2}}$, $\Omega_2=(\alpha'\Sigma_2\alpha)^{\frac{1}{2}}$, $\hat{\Sigma}_1=\frac{1}{T_1-k}\sum_{t=1}^{T_1}\Delta x_t(1-x_{t-1}'(\sum_{t=1}^{T_1}x_{t-1}x_{t-1}')^{-1}x_{t-1})\Delta x_t'$, $\hat{\Sigma}_2=\frac{1}{T-T_1-k}\sum_{t=T_1+1}^{T}\Delta x_t(1-x_{t-1}'(\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}')^{-1}x_{t-1})\Delta x_t'$, $cov(\beta'x)_1=\beta'\sum_{i=0}^{\infty}C_i^*\Sigma_1C_i^{*\prime}\beta$, $cov(\beta'x)_2=\beta'\sum_{i=0}^{\infty}C_i^*\Sigma_2C_i^{*\prime}\beta$, $(\ )_1^{-1}$ *are the first* kr *rows of* $(\ )^{-1}$.

Proof: see appendix.

Theorem 4 shows that the cointegrating vector estimator $\hat{\beta}$ has a normal limiting distribution. When we use a cointegrating vector estimator which neglects the heteroscedasticity of the disturbances, we cannot find accurate expressions of its covariance matrix such that it is hard to test hypotheses on the cointegrating vector in that case. Although the cointegrating relationships are not weakly stationary in this case, as they have a different variance in each of the two variance regimes, they still show mean reversion. The estimators and limiting distributions from theorem 4 can be extended to more variance shifts and other moment conditions (relationships) for the variances can be incorporated. The limiting distribution of the optimal value of the GMM objective function now depends on the relative change of the covariance matrix and the point of change, T_1. As it is not known what the true values of these parameters are, they are typically replaced by sample estimates. The resulting distribution is in that case no longer the true limiting distribution but only an approximation of it. In the next subsection, we will show the applicability of a nonparameteric correction for heteroscedasticity, the use of a White covariance matrix estimator, see [15], in the GMM objective function.

4.2 Cointegration Estimators involving Nonparametric Heteroscedasticity Corrections

For the case of general kind of heteroscedasticity, the White covariance matrix estimator, see [15], can be used in the GMM objective function. This kind of analysis is known as quasi-maximum likelihood or quasi-GMM analysis as we leave part of the stochastic process (conditional variances) unspecified.

We analyze the behavior of the resulting cointegrating vector estimator, using the White covariance matrix estimator, for the case analyzed in the previous subsection, i.e. a change of variance at T_1. The GMM objective function then becomes,

$$G(\alpha,\beta) = vec(\sum_{t=1}^{T} \varepsilon_t x_{t-1}')'(\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1} vec(\sum_{t=1}^{T} \varepsilon_t x_{t-1}'), \qquad (64)$$

where $\hat{\varepsilon}_t$ are the residuals from the unrestricted model estimated assuming homoscedasticity. Theorem 5 states the different cointegration estimators and the limiting distributions of these estimators and the optimal value of the GMM objective function. As the convergence of the White covariance matrix estimator is proved in [15] for the stationary case, a lemma in the

appendix contains a proof of its convergence in the case of nonstationary unit root type series. Note that extensions of the White covariance matrix, like the Newey-West covariance matrix estimator, see [8], which also account for serial correlation, cannot be applied here as neglected serial correlation leads to inconsistent estimators while these covariance estimators can only be applied when consistent estimators are used, see also [10] and [13].

Theorem 5 *When the datagenerating process of the model in equations (56), (57) is such that the number of cointegrating vectors equals r (k-r unit roots), the estimators,*

$$\begin{aligned}\hat{\alpha} &= (\sum_{t=1}^{T} \Delta x_t (1 - x_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x_{2t-1}')^{-1}x_{2t-1})x_{1t-1}') \\ &\quad (\sum_{t=1}^{T} x_{1t-1}(1 - x_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x_{2t-1}')^{-1}x_{2t-1})x_{1t-1}')^{-1}\end{aligned} \tag{65}$$

and

$$\begin{aligned}vec(\hat{\beta}_2') &= -(\sum_{t=1}^{T} x_{t-1}x_{t-1}' \otimes I_r)_2^{-1}((I_k \otimes \hat{\alpha}') \\ &\quad (\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}(I_k \otimes \hat{\alpha}))^{-1} \\ &\quad (I_k \otimes \hat{\alpha}')(\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}vec(\sum_{t=1}^{T}\Delta x_t x_{t-1}') \\ &= vec(\beta_2') - (\sum_{t=1}^{T} x_{t-1}x_{t-1}' \otimes I_r)_2^{-1}((I_k \otimes \hat{\alpha}') \\ &\quad (\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}(I_k \otimes \hat{\alpha}))^{-1}(I_k \otimes \hat{\alpha}') \\ &\quad (\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}vec(\sum_{t=1}^{T}\varepsilon_t x_{t-1}'),\end{aligned} \tag{66}$$

have a limiting behavior which can be characterized by

$$\begin{aligned}\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow\ & n(0, ((wcov(\beta'x)_1 + (1-w)cov(\beta'x)_2) \otimes I_k)^{-1} \\ & (w(cov(\beta'x)_1 \otimes \Sigma_1) + (1-w)(cov(\beta'x)_2 \otimes \Sigma_2)) \\ & ((wcov(\beta'x)_1 + (1-w)cov(\beta'x)_2) \otimes I_k)^{-1}),\end{aligned} \tag{67}$$

and

$$T(vec(\beta_2' - \hat{\beta}_2') \tag{68}$$

$$
\begin{aligned}
\Rightarrow\ & ((\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1 W_1' \Lambda_1' \otimes I_r) + (\int_w^1 (\Lambda_2 W_1(t) \\
& + \Lambda_1 W_1(w))(\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt \otimes I_r))^{-1}[(I_{k-r} \otimes \alpha') \\
& ((\Lambda_1 \int_0^w W_1 W_1' \Lambda_1' \otimes \Sigma_1) + (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))(\Lambda_2 W_1(t) \\
& + \Lambda_1 W_1(w))' dt \otimes \Sigma_2))^{-1}(I_{k-r} \otimes \alpha)]^{-1}(I_{k-r} \otimes \alpha') \\
& ((\Lambda_1 \int_0^w W_1 W_1' \Lambda_1' \otimes \Sigma_1) + (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w)) \\
& (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt \otimes \Sigma_2))^{-1} vec[\Sigma_1^{\frac{1}{2}} \int_0^w dW_2(t)(\Lambda_1 W_1(t))' dt \\
& + \Sigma_2^{\frac{1}{2}} \int_w^1 dW_2(t)(\Lambda_1 W_1(w) + \Lambda_2 W_1(t))' dt]
\end{aligned}
$$

The limiting behavior of the optimal value of the GMM objective function, can be characterized by

$$
\begin{aligned}
& G(\hat{\alpha}, \hat{\beta}) \qquad (69) \\
\Rightarrow\ & vec[\Lambda_1(\int_0^w dW_1 W_1')\Lambda_1' + \Lambda_2(\int_w^1 dW_1(t)(\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt)]' \\
& ((\Lambda_1 \int_0^w W_1 W_1' \Lambda_1' \otimes \alpha_\perp' \Sigma_1^{-1} \alpha_\perp) + (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w)) \\
& (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt \otimes \alpha_\perp' \Sigma_2^{-1} \alpha_\perp))^{-1} vec[\Lambda_1(\int_0^w dW_1 W_1')\Lambda_1' \\
& + \Lambda_2(\int_w^1 dW_1(t)(\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt)]
\end{aligned}
$$

where $w = \frac{T_1}{T}$, W_1 *and* W_2, *are stochastically independent* r, $(k-r)$ *dimensional Brownian motions with identity covariance matrices,* $\Lambda_1 = (\alpha_\perp' \Sigma_1^{-1} \alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2 = (\alpha_\perp' \Sigma_2^{-1} \alpha_\perp)^{\frac{1}{2}}$, $\Omega_1 = (\alpha' \Sigma_1 \alpha)^{\frac{1}{2}}$, $\Omega_2 = (\alpha' \Sigma_2 \alpha)^{\frac{1}{2}}$, $cov(\beta' x)_1 = \beta' \sum_{i=0}^{\infty} C_i^* \Sigma_1 C_i^{*\prime} \beta$, $cov(\beta' x)_2 = \beta' \sum_{i=0}^{\infty} C_i^* \Sigma_2 C_i^{*\prime} \beta$, $(\ )_2^{-1}$ *are the last* $k(k-r)$ *rows of* $(\ )^{-1}$.

Proof: see appendix.

Theorem 5 shows that the limiting behavior of the optimal value of the GMM objective function is identical to the case of specified heteroscedasticity stated in theorem 4. The limiting distributions of the cointegration estimators $\hat{\alpha}$ and $\hat{\beta}$ are, however, different although they are both normal and have a larger variance compared to the limiting distributions of the estimators discussed in theorem 4. As we didnot incorporate any specification of the form of heteroscedasticity, the quasi GMM cointegration estimators discussed previously can also be used in case of several changes in variance and essentially lead to consistent covariances as long as the consistency conditions for the White covariance matrix are fulfilled. The specification of the estimators are identical in that case to the ones in theorem 5 and they also retain their asymptotic normality. The expressions of the asymptotic variances do, however, change.

In this section, an extension of the standard cointegration approach is discussed in the sense that we allow for heteroscedastic disturbances, which extends the results for constant conditional variances. The next section gives another extension to breaks in the cointegrating vector and/or multiplicator.

5 Cointegration with structural breaks

In this section, we investigate the influence of a change in the value of the multiplicator, α, and cointegrating vector, β, at T_1. The model, therefore, is

$$\begin{aligned} \Delta x_t &= \alpha\beta' x_{t-1} + \varepsilon_t \qquad t = 1, ..., T_1, \\ \Delta x_t &= \theta\gamma' x_{t-1} + \varepsilon_t \qquad t = T_1 + 1, ..., T, \end{aligned} \tag{70}$$

where ε_t are Gaussian white noise disturbances with covariance matrix Σ. The GMM objective function corresponding with this model reads,

$$\begin{aligned} G(\alpha, \beta, \gamma, \theta) &= vec(\sum_{t=1}^{T_1} \varepsilon_t x'_{t-1}, \sum_{t=T_1+1}^{T} \varepsilon_t x'_{t-1})' \\ &\begin{pmatrix} ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1})^{-1} \otimes \Sigma^{-1}) & 0 \\ 0 & ((\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1})^{-1} \otimes \Sigma^{-1}) \end{pmatrix} \\ &vec(\sum_{t=1}^{T_1} \varepsilon_t x'_{t-1}, \sum_{t=T_1+1}^{T} \varepsilon_t x'_{t-1}), \end{aligned} \tag{71}$$

where $vec(A, B) = (vec(A)' \; vec(B)')'$. In theorem 6, the cointegration estimators and their limiting distributions are stated jointly with the limiting

distribution of the GMM objective function. As the cointegrating vector estimators and multiplicators all have normal limiting distribution, standard χ^2 tests can be performed to test for the equality of the parameters in each of the two periods. Theorem 6 also states the estimators and their limiting distributions, which can be used when either the cointegrating vectors or multiplicators in each of the two periods are equal to one another.

Theorem 6 *When the DGP in equation (70) is such that the number of cointegrating vectors is r (k-r unit roots), the estimators,*

$$\hat{\alpha} = (\sum_{t=1}^{T_1} \Delta x_t (1 - x'_{2t-1} (\sum_{t=1}^{T_1} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1}) \tag{72}$$
$$(\sum_{t=1}^{T_1} x_{1t-1} (1 - x'_{2t-1} (\sum_{t=1}^{T_1} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1})^{-1},$$

$$\hat{\theta} = (\sum_{t=T_1+1}^{T} \Delta x_t (1 - x'_{2t-1} (\sum_{t=T_1+1}^{T} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1}) \tag{73}$$
$$(\sum_{t=T_1+1}^{T} x_{1t-1} (1 - x'_{2t-1} (\sum_{t=T_1+1}^{T} x_{2t-1} x'_{2t-1})^{-1} x_{2t-1}) x'_{1t-1})^{-1},$$

and

$$\hat{\beta} = (\sum_{t=1}^{T_1} x_{t-1} x'_{t-1})^{-1} (\sum_{t=1}^{T_1} x_{t-1} \Delta x'_t) \Sigma_1^{-1} \hat{\alpha} (\hat{\alpha}' \Sigma_1^{-1} \hat{\alpha})^{-1}, \tag{74}$$

$$\hat{\gamma} = (\sum_{t=T_1+1}^{T} x_{t-1} x'_{t-1})^{-1} (\sum_{t=T_1+1}^{T} x_{t-1} \Delta x'_t) \Sigma_2^{-1} \hat{\theta} (\hat{\theta}' \Sigma_2^{-1} \hat{\theta})^{-1}, \tag{75}$$

have a limiting behavior which can be characterized by,

$$\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow n(0,\ cov(\beta' x)^{-1} \otimes w\Sigma), \tag{76}$$

$$\sqrt{T}(\hat{\theta} - \theta) \Rightarrow n(0,\ cov(\gamma' x)^{-1} \otimes (1-w)\Sigma), \tag{77}$$

and

$$T(\hat{\beta} - \beta) \Rightarrow \begin{pmatrix} 0 \\ (\beta'_\perp \beta_\perp)^{-1} \beta'_\perp \alpha_\perp \Lambda_1^{-1\prime} (\int_0^w W_1 W_1')^{-1} (\int_0^w W_1 dW_2') \Omega_1' \end{pmatrix}, \tag{78}$$

$$T(\gamma_2 - \hat{\gamma}_2) \Rightarrow [(\int_w^1 (\gamma'_\perp \beta_\perp (\alpha'_\perp \beta_\perp)^{-1} \Lambda_1 W_1(w) + \gamma'_\perp \gamma_\perp (\theta'_\perp \gamma_\perp)^{-1} \Lambda_2 W_1(t)) \tag{79}$$

$$(\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w)+\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))'dt)]^{-1}$$
$$\int_w^1[\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w)+\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))$$
$$dW_2(t)'dt]\Omega_2'.$$

The limiting behavior of the optimal value of the objective function can be characterized by,

$$G(\hat\alpha,\hat\beta,\hat\gamma,\hat\theta) \quad (80)$$
$$\Rightarrow vec((\int_0^w W_1 dW_1')')'((\int_0^w W_1W_1')^{-1}\otimes I_{k-r})vec((\int_0^w W_1dW_1')')$$
$$+vec(\int_w^1[\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1W_1(w)$$
$$+W_1(t)]dW_1(t)')'([\int_w^1(\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1W_1(w)$$
$$+W_1(t))(\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1W_1(w)$$
$$+W_1(t))'dt]^{-1}\otimes I_{k-r})vec(\int_w^1[\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1W_1(w)$$
$$+W_1(t)]dW_1(t)').$$

When the model in equation (70) is such that the cointegrating vectors are equal in the two periods, $\beta=\gamma$, which can be tested for using a χ^2 test, the GMM estimator for β reads (estimators for α and θ result from the first part),

$$vec(\hat\beta')=[(\sum_{t=1}^{T_1}x_{t-1}x_{t-1}'\otimes\hat\alpha'\Sigma_1^{-1}\hat\alpha))+ \quad (81)$$
$$(\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}'\otimes\hat\theta'\Sigma_2^{-1}\hat\theta)]^{-1}[(I_k\otimes\hat\alpha'\Sigma_1^{-1})$$
$$(vec(\sum_{t=1}^{T_1}\Delta x_t x_{t-1}')+(I_k\otimes\hat\theta'\Sigma_2^{-1})(vec(\sum_{t=T_1+1}^{T}\Delta x_t x_{t-1}'))]$$
$$=vec(\begin{pmatrix}I_r\\-\beta_2\end{pmatrix}')+[(\sum_{t=1}^{T_1}x_{t-1}x_{t-1}'\otimes\hat\alpha'\Sigma_1^{-1}\hat\alpha))$$
$$+(\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}'\otimes\hat\theta'\Sigma_2^{-1}\hat\theta)]^{-1}[vec(\sum_{t=1}^{T_1}\hat\alpha'\Sigma_1^{-1}\varepsilon_t x_{t-1}')$$

$$+vec(\sum_{t=T_1+1}^{T}\hat{\theta}'\Sigma_2^{-1}\varepsilon_t x'_{t-1})]$$

and the limiting behavior of this estimator can be characterized by,

$$T vec(\hat{\beta}_2-\beta_2) \quad (82)$$

$$\Rightarrow((\beta'_\perp\beta_\perp)^{-1}\otimes I_r)[((\alpha'_\perp\beta_\perp)^{-1}\Lambda_1(\int_0^w W_1W'_1)\Lambda'_1(\beta'_\perp\alpha_\perp)^{-1}\otimes\alpha'\Sigma^{-1}\alpha)$$

$$+([\int_w^1((\alpha'_\perp\beta_\perp)^{-1}\Lambda_1W_1(w)+(\theta'_\perp\beta_\perp)^{-1}\Lambda_2W_1(t))((\alpha'_\perp\beta_\perp)^{-1}\Lambda_1W_1(w)$$

$$+(\theta'_\perp\beta_\perp)^{-1}\Lambda_2W_1(t))'dt]\otimes\theta'\Sigma^{-1}\theta)]^{-1}vec(\Omega_1(\int_0^w dW_2W'_1)\Lambda'_1(\beta'_\perp\alpha_\perp)^{-1}$$

$$+\Omega_2\int_w^1[dW_2(t)(W'_1(w)\Lambda'_1(\beta'_\perp\alpha_\perp)^{-1}+W'_1(t)\Lambda'_2(\beta'_\perp\theta_\perp)^{-1})dt]).$$

When the model in equation (70) is such that the multiplicators of cointegrating vectors are equal in the two periods, $\alpha=\theta$, which can be tested for using a χ^2 test, the GMM estimator for α reads (estimators for β and γ result from the first part)

$$\hat{\alpha}=(\sum_{t=1}^{T_1}\Delta x_t x'_{t-1}\hat{\beta}+\sum_{t=T_1+1}^{T}\Delta x_t x'_{t-1}\hat{\gamma})^{-1}(\hat{\beta}'\sum_{t=1}^{T_1}x_{t-1}x'_{t-1}\hat{\beta}+\hat{\gamma}'\sum_{t=T_1+1}^{T}x_{t-1}x'_{t-1}\hat{\gamma})^{-1} \quad (83)$$

and its limiting behavior can be characterized by

$$\sqrt{T}(\hat{\alpha}-\alpha)\Rightarrow n(0,wcov(\beta'x)_1+(1-w)cov(\gamma'x)_2) \quad (84)$$

where $w=\frac{T_1}{T}$, W_1 and W_2, are stochastically independent r, $(k-r)$ dimensional Brownian motions with identity covariance matrices, $\Lambda_1=(\alpha'_\perp\Sigma^{-1}\alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2=(\theta'_\perp\Sigma^{-1}\theta_\perp)^{\frac{1}{2}}$, $\Omega_1=(\alpha'\Sigma\alpha)^{\frac{1}{2}}$, $\Omega_2=(\theta'\Sigma\theta)^{\frac{1}{2}}$, $\hat{\Sigma}_1=\frac{1}{T_1-k}\sum_{t=1}^{T_1}\Delta x_t(1-x'_{t-1}(\sum_{t=1}^{T_1}x_{t-1}x'_{t-1})^{-1}x_{t-1})\Delta x'_t$, $\hat{\Sigma}_2=\frac{1}{T-T_1-k}\sum_{t=T_1+1}^{T}\Delta x_t(1-x'_{t-1}(\sum_{t=T_1+1}^{T}x_{t-1}x'_{t-1})^{-1}x_{t-1})\Delta x'_t$, $cov(\beta'x)_1=\beta'\sum_{i=0}^{\infty}C^*_{1i}\Sigma_1C^{*\prime}_{1i}\beta$, $cov(\gamma'x)_2=\beta'\sum_{i=0}^{\infty}C^*_{2i}\Sigma_2C^{*\prime}_{2i}\beta$, and $C_1(L)$, $C_2(L)$ are the Vector Moving Average representations of the first and second subsets.

Proof: see appendix.

Theorem 6 again shows that the GMM estimators of the cointegrating vector and multiplicator have normal limiting distributions in case of breaks in the cointegrating vector and/or multiplicator. Similar to the limiting distribution of the optimal value of the GMM objective function in case of heteroscedasticity, the limiting distribution of the optimal value of the GMM objective function again depends on model parameters and the changing point T_1. An approximation of this limiting distribution can again be constructed using the estimated values of the parameters, α, β, θ, γ and T_1. As this leads to a rather complicated testing procedure, it may be preferable to fix the number of cointegrating vectors a priori and just perform tests on the estimated cointegrated vectors and multiplicators, which are straightforward to construct. This reasoning also holds for the cointegration tests discussed in the previous section.

6 Conclusions

A GMM framework for cointegration analysis is developed allowing for extensions of the models, which are analyzable using the methods documented in the literature. As examples, extensions along the lines of heteroscedasticity and structural breaks are included and the resulting cointegration estimators are shown to have normal limiting distributions while the optimal value of the GMM objective function has a limiting distribution, which is a Brownian motion functional with additional parameters resulting from the change of properties of the involved Brownian motions. These additional parameters are essentially the parameters in the model with vary over time resulting in heteroscedasticity or structural breaks. In future work, we will apply the developed framework for a.o. cointegration analysis in financial series, for example term structure of interest rates. As heteroscedasticity is a stylized fact of these series, the standard cointegration procedures cannot be applied here as they lead to inconsistent estimators and/or incorrect (asymptotic) variances of the estimators.

Appendix

Lemma 1.

In this lemma the consistency of the White Covariance Matrix Estimator is proved for the case of nonstationary cointegrated regressors. The proof is given for the homoscedastic case, extensions to heteroscedasticity follow naturally as a homoscedastic dataset can be interpreted as a subset of a heteroscedastic dataset. It is assumed that $v_t = \hat{\varepsilon}_t\hat{\varepsilon}_t' - \Sigma$, $vec(v_t) \sim f(0, \Lambda \otimes I_k)$, $E(x_{t-1}x_{t-1}' \otimes v_t | I_t) = 0$. We assume a DGP of the form,

$$\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t,$$

To proof consistency of the White Covariance matrix estimator $\sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t')$, we first analyze its behavior in terms of the cointegrating relationships and their orthogonal complements,

$$\begin{aligned}
& \left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right)' \sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t')\left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right) \\
= & \left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right)' \sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \Sigma + v_t)\left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right) \\
= & \left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right)' \sum_{t=1}^{T}(x_{t-1}x_{t-1}' \otimes \Sigma)\left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k\right) \\
& + \left(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix}\right)' \sum_{t=1}^{T} x_{t-1}x_{t-1}' \begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes v_t)
\end{aligned}$$

The cointegrating relationships are stationary such that the standard results apply to them,

$$\frac{1}{T}(\sum_{t=1}^{T} \beta' x_{t-1}x_{t-1}'\beta \otimes v_t) \Rightarrow 0.$$

For the nonstationary case it holds according to the central limit theorem, that because x_{t-1} and v_t are uncorrelated,

$$\begin{aligned}
& \frac{1}{T^{3/2}}(\sum_{t=1}^{T} \beta_\perp' x_{t-1}x_{t-1}'\beta_\perp \otimes v_t)_{ij} \\
= & (\frac{1}{T^{3/2}} \sum_{t=1}^{T} \beta_{i\perp}' x_{t-1}x_{t-1}'\beta_{j\perp} v_t) \\
\Rightarrow & \; n(0, \Lambda \otimes c(\int W_i^2 W_j^2) I_k).
\end{aligned}$$

where c is a function of the cointegrating vectors, multiplicators and covariance Σ, W_i and W_j are (possibly) correlated Brownian Motions. So,

$$\frac{1}{T^2}(\sum_{t=1}^{T} \beta_\perp' x_{t-1} x_{t-1}' \beta_\perp \otimes v_t) \Rightarrow 0$$

This result can also be applied to the cross products of the cointegrating relationships and their orthogonal complements. Consequently,

$$\begin{aligned} & (\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k)' \sum_{t=1}^{T} (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t \hat{\varepsilon}_t')(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k) \\ \Rightarrow \quad & (\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k)' \sum_{t=1}^{T} (x_{t-1}x_{t-1}' \otimes \Sigma)(\begin{pmatrix} \frac{1}{\sqrt{T}}\beta & \frac{1}{T}\beta_\perp \end{pmatrix} \otimes I_k) \end{aligned}$$

which proofs that the White Covariance Matrix estimator can be used in the case of a cointegrated dataset.

Proof of theorem 1.

In [4] it is proved that the stochastic process x_t, from equation (1), can be represented by

$\Delta x_t = C(L)\Sigma^{\frac{1}{2}}\xi_t$, $\beta = \begin{pmatrix} I_r & -\beta_2' \end{pmatrix}'$

where ξ_t is a $k-$variate Gaussian white noise process with zero mean and identity covariance matrix. Consequently,

$$x_t = \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'\Sigma^{\frac{1}{2}} \sum_{j=1}^{t} \xi_j + C^*(L)\Sigma^{\frac{1}{2}}\xi_t,$$

$$x_{1t} = \beta_2'(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'\Sigma^{\frac{1}{2}} \sum_{j=1}^{t} \xi_j + \begin{pmatrix} I_r \\ 0 \end{pmatrix} C^*(L)\Sigma^{\frac{1}{2}}\xi_t,$$

$$x_{2t} = (\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'\Sigma^{\frac{1}{2}} \sum_{j=1}^{t} \xi_j + \begin{pmatrix} 0 \\ I_{k-r} \end{pmatrix} C^*(L)\Sigma^{\frac{1}{2}}\xi_t,$$

where $C(L) = C(1) + (1-L)C^*(L)$, $C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i$. The least squares estimator of α, $\hat{\alpha}$, can also be expressed as

$$\begin{aligned} \hat{\alpha} - \alpha &= (\sum_{t=1}^{T} u_t(1 - x_{2t-1}'(\sum_{t=1}^{T} x_{2t-1}x_{2t-1}')^{-1}x_{2t-1})x_{1t-1}') \\ & \quad (\sum_{t=1}^{T} x_{1t-1}(1 - x_{2t-1}'(\sum_{t=1}^{T} x_{2t-1}x_{2t-1}')^{-1}x_{2t-1})x_{1t-1}')^{-1} \\ &= (\sum_{t=1}^{T} u_t(x_{1t-1} - \tilde{\beta}_2' x_{2t-1})')(\sum_{t=1}^{T} (x_{1t-1} - \tilde{\beta}_2' x_{2t-1})(x_{1t-1} - \tilde{\beta}_2' x_{2t-1})')^{-1} \end{aligned}$$

with $\tilde{\beta}_2 = (\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1}x'_{1t-1}$. $\tilde{\beta}_2$ is a superconsistent estimator of β_2 and can therefore be treated as equal to β_2 in the derivation of the limiting distribution of $\hat{\alpha}$. Since

$$\frac{1}{T}\sum_{t=1}^{T}(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})' \Rightarrow cov(\beta' x) = \beta' \sum_{i=0}^{\infty} C_i^* \Sigma C_i^{*\prime} \beta,$$

and, $T^{-\frac{1}{2}}(\sum_{t=1}^{T} u_t(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})') \Rightarrow n(0,\ cov(\beta' x) \otimes \Sigma)$, the limiting distribution of $\hat{\alpha}$ becomes

$$\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow n(0,\ cov(\beta' x)^{-1} \otimes \Sigma).$$

With respect to the cointegrating vector,

$$\begin{aligned}
\hat{\beta} &= (\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{T} x_{t-1}\Delta x'_t)\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} I_r \\ (\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}(\sum_{t=1}^{T} x_{t-1}(x'_{t-1}\beta\alpha' + u'_t))\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \end{pmatrix} \\
&= \begin{pmatrix} I_r \\ (\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}(\sum_{t=1}^{T} x_{t-1}(x'_{t-1}\beta\alpha' + u'_t))\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} \end{pmatrix} \\
&= \begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix} + \begin{pmatrix} 0 \\ (\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}(\sum_{t=1}^{T} x_{t-1}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} \end{pmatrix}
\end{aligned}$$

where $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}$ indicates the last $(k-r)$ rows of $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}$ and $\hat{\alpha}$ is a consistent estimator of α such that the difference between $\hat{\alpha}$ and α will only affects orders of convergence exceeding T. Furthermore $\hat{\alpha} = (\sum_{t=1}^{T} x_{t-1}x'_{t-1})_1^{-1}(\sum_{t=1}^{T} x_{t-1}\Delta x'_t)$, where $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})_1^{-1}$ indicates the first r rows of $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}$.

To analyze the limiting behavior of $\hat{\beta}$, we have to determine the limiting expressions of both $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}$ and $(\sum_{t=1}^{T} x_{t-1}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1}$. Starting with the latter expression, its limiting behavior can be analyzed using the stochastic trend specification.

$$\begin{aligned}
(\sum_{t=1}^{T} x_{t-1}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} &= (\sum_{t=1}^{T} \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp\Sigma^{\frac{1}{2}}(\sum_{j=1}^{t-1}\xi_j)\xi'_t\Sigma^{-\frac{1}{2}}\alpha \\
&\quad + \sum_{t=1}^{T} C^*(L)\Sigma^{\frac{1}{2}}\xi_{t-1}\xi'_t\Sigma^{-\frac{1}{2}}\alpha)(\alpha'\Sigma^{-1}\alpha)^{-1}
\end{aligned}$$

Since $\Sigma^{\frac{1}{2}}\alpha_\perp$ is orthogonal to $\Sigma^{-\frac{1}{2}}\alpha$, i.e. $(\Sigma^{\frac{1}{2}}\alpha_\perp)'\Sigma^{-\frac{1}{2}}\alpha = \alpha_\perp'\alpha = 0$, the brownian motions appearing in the limiting expression are independent,

$$\frac{1}{T}\sum_{t=1}^{T}\alpha_\perp'\Sigma^{\frac{1}{2}}(\sum_{j=1}^{t-1}\xi_j)\xi_t'\Sigma^{-\frac{1}{2}}\alpha \Rightarrow \Lambda_1\int W_1 dW_2'\Lambda_2',$$

since $\frac{1}{\sqrt{T}}\alpha_\perp'\Sigma^{\frac{1}{2}}(\sum_{j=1}^{T-1}\xi_j) \Rightarrow \Lambda_1 W_1$, W_1 is a $(k-r)$ dimensional brownian motion with covariance matrix I_{k-r} and $\Lambda_1 = (\alpha_\perp'\Sigma\alpha_\perp)^{\frac{1}{2}}$, W_2 is a r dimensional brownian motion with covariance matrix I_r and W_2 is stochastically independent of W_1, $\Lambda_2 = (\alpha'\Sigma^{-1}\alpha)^{\frac{1}{2}}$.

Also the limiting behavior of $(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1}$ is determined by the stochastic trend specification.

$$(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} = \begin{pmatrix} \beta & \beta_\perp \end{pmatrix}[\begin{pmatrix} \beta & \beta_\perp \end{pmatrix}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}')\begin{pmatrix} \beta & \beta_\perp \end{pmatrix}]^{-1}\begin{pmatrix} \beta & \beta_\perp \end{pmatrix}'$$

So, the limiting behavior of $\begin{pmatrix} \beta & \beta_\perp \end{pmatrix}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}')\begin{pmatrix} \beta & \beta_\perp \end{pmatrix}$ is of primary importance.

$$\begin{pmatrix} T^{-\frac{1}{2}}\beta & T^{-1}\beta_\perp \end{pmatrix}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}')\begin{pmatrix} T^{-\frac{1}{2}}\beta & T^{-1}\beta_\perp \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} cov(\beta'x) & 0 \\ 0 & \beta_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1(\int W_1W_1')\Lambda_1'(\alpha_\perp'\beta_\perp)^{-1\prime}\beta_\perp'\beta_\perp \end{pmatrix}$$

as

$$T^{-1}\sum_{t=1}^{T}\beta'C^*(L)\Sigma^{\frac{1}{2}}\xi_{t-1}\xi_{t-1}'\Sigma^{\frac{1}{2}\prime}C^*(L)\beta \Rightarrow cov(\beta'x),$$

$$T^{-2}\beta_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\sum_{t=1}^{T}(\alpha_\perp'\Sigma^{\frac{1}{2}}\sum_{j=1}^{t-1}\xi_j)(\sum_{j=1}^{t-1}\xi_j'\Sigma^{\frac{1}{2}\prime}\alpha_\perp)(\alpha_\perp'\beta_\perp)^{-1\prime}\beta_\perp'\beta_\perp$$

$$\Rightarrow \beta_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1(\int W_1W_1')\Lambda_1'(\alpha_\perp'\beta_\perp)^{-1\prime}\beta_\perp'\beta_\perp,$$

$$T^{-1\frac{1}{2}}\sum_{t=1}^{T}\beta'C^*(L)\Sigma^{\frac{1}{2}}\xi_{t-1}(\sum_{j=1}^{t-1}\xi_j'\Sigma^{\frac{1}{2}\prime}\alpha_\perp)(\alpha_\perp'\beta_\perp)^{-1\prime}\beta_\perp'\beta_\perp \Rightarrow 0.$$

Consequently,

$$[\begin{pmatrix} T^{-\frac{1}{2}}\beta & T^{-1}\beta_\perp \end{pmatrix}'(\sum_{t=1}^{T} x_{t-1}x_{t-1}')\begin{pmatrix} T^{-\frac{1}{2}}\beta & T^{-1}\beta_\perp \end{pmatrix}]^{-1}$$

$$\Rightarrow \begin{pmatrix} cov(\beta'x)^{-1} & 0 \\ 0 & (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W_1')^{-1}\Lambda_1^{-1}\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1} \end{pmatrix}$$

and

$$
\begin{aligned}
& (\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1} \\
\Rightarrow \quad & O(T)\beta cov(\beta' x)^{-1}\beta' + \\
& O(T^2)\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W_1')^{-1}\Lambda_1^{-1}\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp.
\end{aligned}
$$

where $O(T^j)$ indicates that the limiting behavior of this part is proportional to T^j. The latter part governs the limiting behavior of $(\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1}$, which can be characterized by

$$
T^2(\sum_{t=1}^{T} x_{t-1}x'_{t-1})_2^{-1} \Rightarrow (\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W_1')^{-1}\Lambda_1^{-1}\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp
$$

as $\beta_\perp = \begin{pmatrix} \beta'_2 \\ I_{k-r} \end{pmatrix}$. So, the limiting expression for the cointegrating vector estimator becomes,

$$
\begin{aligned}
T(\hat{\beta}-\beta) \quad & \Rightarrow \quad \begin{pmatrix} 0 \\ (\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W_1')^{-1}(\int W_1dW_2')\Lambda_2' \end{pmatrix} \\
& \Rightarrow \quad \begin{pmatrix} 0 \\ n(0, \alpha'\Sigma^{-1}\alpha \otimes \Theta \end{pmatrix},
\end{aligned}
$$

where $\Theta = (\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W_1')^{-1}\Lambda_1^{-1}\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}$ and can be approximated by $(\frac{1}{T^2}\sum_{t=1}^{T} x_{2t-1}(1 - x'_{1t-1}(\sum_{t=1}^{T} x_{1t-1}x'_{1t-1})^{-1}x_{1t-1})x'_{2t-1})^{-1}$ ($\Rightarrow \Theta$).

Proof of theorem 2 (only the second part of theorem 2 is proved).

When the DGP of x_t reads,

$$
\Delta x_t = \alpha_\perp\lambda' + \alpha(\beta' x_{t-1} + \mu') + \varepsilon_t,
$$

$c = \alpha_\perp\lambda' + \alpha\mu'$, it has the stochastic trend representation, see [4],

$\Delta x_t = C(L)(c + \Sigma^{\frac{1}{2}}\xi_t)$, $\beta = \begin{pmatrix} I_r & -\beta'_2 \end{pmatrix}'$,

where ξ_t is a $k-$variate Gaussian white noise process with zero mean and identity covariance matrix. Consequently,

$$
x_t = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp(t\alpha_\perp\lambda' + \Sigma^{\frac{1}{2}}\sum_{j=1}^{t}\xi_j) + C^*(1)\alpha\mu' + C^*(L)\Sigma^{\frac{1}{2}}\xi_t,
$$

$$x_{1t} = \beta_2'(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'(t\alpha_\perp\lambda' + \Sigma^{\frac{1}{2}}\sum_{j=1}^{t}\xi_j) + \begin{pmatrix} I_r \\ 0 \end{pmatrix}(C^*(1)\alpha\mu' + C^*(L)\Sigma^{\frac{1}{2}}\xi_t),$$

$$x_{2t} = (\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'(t\alpha_\perp\lambda' + \Sigma^{\frac{1}{2}}\sum_{j=1}^{t}\xi_j) + \begin{pmatrix} 0 \\ I_{k-r} \end{pmatrix}(C^*(1)\alpha\mu' + C^*(L)\Sigma^{\frac{1}{2}}\xi_t),$$

where $C(L) = C(1) + (1-L)C^*(L)$, $C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i$, $\beta' C^*(1)\alpha = I_r$.

The least squares estimator of α, $\hat{\alpha}$, can also be expressed as

$$\begin{aligned}
&\hat{\alpha} - \alpha = \\
&(\sum_{t=1}^{T} u_t(1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}'(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}') \\
&(\sum_{t=1}^{T} x_{1t-1}(1 - \begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}'(\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})x_{1t-1}')^{-1} \\
= \quad &(\sum_{t=1}^{T} u_t(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}'\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})')(\sum_{t=1}^{T}(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}' \\
&\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}'\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}))^{-1}
\end{aligned}$$

with $\begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix} = (\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}')^{-1}\sum_{t=1}^{T}\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix}x_{1t-1}'$. $\tilde{\beta}_2$ is a superconsistent estimator of β_2. Since

$\frac{1}{T}\sum_{t=1}^{T}(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}'\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}'\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})'$

$\Rightarrow cov(\beta' x - \mu') = \beta' \sum_{i=0}^{\infty} C_i^* \Sigma C_i^{*\prime}\beta$, and,

$T^{-\frac{1}{2}}(\sum_{t=1}^{T} u_t(x_{1t-1} - \begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\mu} \end{pmatrix}'\begin{pmatrix} x_{2t-1} \\ 1 \end{pmatrix})') \Rightarrow n(0,\ cov(\beta' x - \mu') \otimes \Sigma)$, the limiting distribution of $\hat{\alpha}$ becomes

$$\sqrt{T}(\hat{\alpha} - \alpha) \Rightarrow n(0,\ cov(\beta' x - \mu')^{-1} \otimes \Sigma).$$

With respect to the cointegrating vector,

$$\begin{aligned}
\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \end{pmatrix} &= (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1}(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\Delta x_t')\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} I_r \\ -\hat{\beta}_2 \\ \hat{\mu} \end{pmatrix}
\end{aligned}$$

$$\begin{pmatrix} -\hat{\beta}_2 \\ \hat{\mu} \end{pmatrix} = ((\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1}(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}(\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}'\begin{pmatrix} \beta \\ \mu \end{pmatrix}\alpha'$$

$$
\begin{aligned}
& +\lambda\alpha'_{\perp} + u'_t))\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \\
= \; & (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1}(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \\
& (\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}'\begin{pmatrix} \beta \\ \mu \end{pmatrix}\alpha' + u'_t))\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} \\
= \; & \begin{pmatrix} -\beta_2 \\ \mu \end{pmatrix} + (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1} \\
& (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1}
\end{aligned}
$$

as $\alpha'_{\perp}\Sigma^{-1}\alpha = 0$ since $\Sigma^{-1} = P\Lambda P'$, Λ =diag$(\lambda_i) = \sum_{i=1}^{k}\lambda_i e_i e'_i$, $PP' = I_k$, $\alpha'_{\perp}PP'\alpha = b'_{\perp}b = 0$, $P'\alpha = b = (b'_1...b'_k)'$, $\alpha'_{\perp}\Sigma^{-1}\alpha = b'_{\perp}\Lambda b = \sum_{i=1}^{k}\lambda_i b'_{\perp i}b_i = 0$ as $b'_{\perp i}b_i = 0\ \forall i$. $(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1}$ indicates the last $(k - r + 1)$ rows of $(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1}$ and $\hat{\alpha}$ is a consistent estimator of α such that the difference between $\hat{\alpha}$ and α will only affect orders of convergence exceeding T. To analyze the limiting behavior of $\hat{\beta}$, we have to determine the limiting expressions of both $(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1}$ and $(\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1}$. Starting with the latter expression, its limiting behavior can be analyzed using the stochastic trend specification.

$$
\begin{aligned}
& (\sum_{t=1}^{T}\begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}u'_t)\Sigma^{-1}\alpha(\alpha'\Sigma^{-1}\alpha)^{-1} \\
= \; & (\sum_{t=1}^{T}\begin{pmatrix} \beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}\alpha'_{\perp}(t\alpha_{\perp}\lambda' + \Sigma^{\frac{1}{2}}(\sum_{j=1}^{t-1}\xi_j) \\ 1 \end{pmatrix}\xi'_t\Sigma^{-\frac{1}{2}}\alpha \\
& +\sum_{t=1}^{T}\begin{pmatrix} C^*(1)\alpha\mu' + C^*(L)\Sigma^{\frac{1}{2}}\xi_{t-1}) \\ 0 \end{pmatrix}\xi'_t\Sigma^{-\frac{1}{2}}\alpha)(\alpha'\Sigma^{-1}\alpha)^{-1}
\end{aligned}
$$

Since $\Sigma^{\frac{1}{2}}\alpha_{\perp}$ is orthogonal to $\Sigma^{-\frac{1}{2}}\alpha$, i.e. $(\Sigma^{\frac{1}{2}}\alpha_{\perp})'\Sigma^{-\frac{1}{2}}\alpha = \alpha'_{\perp}\alpha = 0$, the brownian motions appearing in the limiting expression are independent,

$$
\begin{pmatrix} T^{-1}\lambda_{\perp} \\ T^{-\frac{3}{2}}\lambda \end{pmatrix}(\alpha'_{\perp}\alpha_{\perp})^{-1}\alpha'_{\perp}\sum_{t=1}^{T}(t\alpha_{\perp}\lambda' + \Sigma^{\frac{1}{2}}(\sum_{j=1}^{t-1}\xi_j))\xi'_t\Sigma^{-\frac{1}{2}}\alpha)
$$

$$\Rightarrow \quad \Lambda_3 \int \begin{pmatrix} W_{11} \\ \tau \end{pmatrix} dW_2' \Lambda_2'$$

and

$$\begin{pmatrix} T^{-1} I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix} \beta_\perp^{*\prime} \quad (\sum_{t=1}^{T} \beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \alpha_\perp' (t \alpha_\perp \lambda' + \Sigma^{\frac{1}{2}} (\sum_{j=1}^{t-1} \xi_j) \xi_t' \Sigma^{-\frac{1}{2}} \alpha)$$

$$\Rightarrow \qquad \Lambda_3 \int \begin{pmatrix} W_{11} \\ \tau \end{pmatrix} dW_2' \Lambda_2'$$

So

$$\begin{pmatrix} \beta_\perp^* \begin{pmatrix} T^{-1} I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix} & 0 \\ 0 & T^{-\frac{1}{2}} \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} u_t') \Sigma^{-1} \alpha (\alpha' \Sigma^{-1} \alpha)^{-1}$$

$$\Rightarrow \quad \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix} \int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} dW_2' \Lambda_2'$$

where $\beta_\perp^* = \beta_\perp (\beta_\perp' \beta_\perp)^{-1} \beta_\perp' \alpha_\perp (\alpha_\perp' \alpha_\perp)^{-1} \begin{pmatrix} \lambda_\perp \\ \lambda \end{pmatrix}'$, $\lambda_\perp \lambda' = 0$, W_{11} is a $(k-r-1)$ dimensional brownian motion with covariance matrix I_{k-r-1} and $\Lambda_3 = (\begin{pmatrix} \lambda_\perp (\alpha_\perp' \alpha_\perp)^{-1} \alpha_\perp' \Sigma^{\frac{1}{2}} \\ \lambda \lambda' \end{pmatrix} \begin{pmatrix} \lambda_\perp (\alpha_\perp' \alpha_\perp)^{-1} \alpha_\perp' \Sigma^{\frac{1}{2}} \\ \lambda \lambda' \end{pmatrix}')^{\frac{1}{2}}$, W_2 is a r dimensional brownian motion with covariance matrix I_r and W_2 is stochastically independent of W_{11}, $\Lambda_2 = (\alpha' \Sigma^{-1} \alpha)^{\frac{1}{2}}$, $\tau(t) = t$, $\iota(t) = 1$, $0 \le t \le 1$.

Also the limiting behavior of $(\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1}$ is determined by the stochastic trend specification.

$$\begin{aligned} & (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')^{-1} \\ = \quad & \begin{pmatrix} \beta & \beta_\perp^* & 0 \\ 0 & 0 & 1 \end{pmatrix} [\begin{pmatrix} \beta & \beta_\perp^* & 0 \\ 0 & 0 & 1 \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}') \\ & \begin{pmatrix} \beta & \beta_\perp^* & 0 \\ 0 & 0 & 1 \end{pmatrix}]^{-1} \begin{pmatrix} \beta & \beta_\perp^* & 0 \\ 0 & 0 & 1 \end{pmatrix}' \end{aligned}$$

$$\begin{pmatrix} T^{-\frac{1}{2}} \beta & \beta_\perp^* \begin{pmatrix} T^{-1} I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix} & 0 \\ 0 & 0 & T^{-\frac{1}{2}} \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')$$

$$\begin{pmatrix} T^{-\frac{1}{2}}\beta & \beta_\perp^* \begin{pmatrix} T^{-1}I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix} & 0 \\ 0 & 0 & T^{-\frac{1}{2}} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} cov(\beta' x - \mu') + \mu'\mu & 0 & \mu' \\ 0 & \Lambda_3(\int \begin{pmatrix} W_{11} \\ \tau \end{pmatrix} \begin{pmatrix} W_{11} \\ \tau \end{pmatrix}')\Lambda_3' & 0 \\ \mu & 0 & 1 \end{pmatrix}$$

Consequently,

$$(\begin{pmatrix} T^{-\frac{1}{2}}\beta & \begin{pmatrix} T^{-1}I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix}\beta_\perp^* & 0 \\ 0 & 0 & T^{-\frac{1}{2}} \end{pmatrix}' (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')$$

$$\begin{pmatrix} T^{-\frac{1}{2}}\beta & \begin{pmatrix} T^{-1}I_{k-r-1} & 0 \\ 0 & T^{-\frac{3}{2}} \end{pmatrix}\beta_\perp^* & 0 \\ 0 & 0 & T^{-\frac{1}{2}} \end{pmatrix})^{-1}$$

$$\Rightarrow \begin{pmatrix} cov(\beta' x - \mu')^{-1} & 0 & -cov(\beta' x - \mu')^{-1}\mu' \\ 0 & \Lambda_3^{-1\prime}(\int \begin{pmatrix} W_{11} \\ \tau \end{pmatrix} \begin{pmatrix} W_{11} \\ \tau \end{pmatrix}')^{-1}\Lambda_3^{-1} & 0 \\ -\mu cov(\beta' x - \mu')^{-1} & 0 & 1 + \mu cov(\beta' x - \mu')^{-1}\mu' \end{pmatrix}$$

The limiting behavior of $(\sum_{t=1}^{T} x_{t-1}x_{t-1}')_2^{-1}$ now becomes

$$\begin{pmatrix} \beta_{2\perp}^* \begin{pmatrix} TI_{k-r-1} & 0 \\ 0 & T^{\frac{3}{2}} \end{pmatrix} & 0 \\ 0 & T^{\frac{1}{2}} \end{pmatrix} (\sum_{t=1}^{T} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix}')_2^{-1}$$

$$\begin{pmatrix} \beta_\perp^* \begin{pmatrix} TI_{k-r-1} & 0 \\ 0 & T^{\frac{3}{2}} \end{pmatrix} & 0 \\ 0 & T^{\frac{1}{2}} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1\prime} (\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}')^{-1} \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

where $\beta_{2\perp}^* = (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp(\alpha_\perp'\alpha_\perp)^{-1} \begin{pmatrix} \lambda_\perp \\ \lambda \end{pmatrix}'$. So, the limiting expression for the cointegrating vector estimator becomes,

$$\begin{pmatrix} \begin{pmatrix} TI_{k-r-1} & 0 \\ 0 & T^{\frac{3}{2}} \end{pmatrix} (\beta_{2\perp}^{*\prime}\beta_{2\perp}^*)^{-1}\beta_{2\perp}^{*\prime} & 0 \\ 0 & T^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \beta_2 - \hat{\beta}_2 \\ \hat{\mu} - \mu \end{pmatrix}$$

$$\Rightarrow \quad \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1\prime} (\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}')^{-1} \int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} dW_2' \Lambda_2'$$

$$\Rightarrow \quad n(0,\ \alpha'\Sigma^{-1}\alpha \otimes \Theta_2)$$

where $\Theta_2 = \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1\prime} (\int \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix} \begin{pmatrix} W_{11} \\ \tau \\ \iota \end{pmatrix}')^{-1} \begin{pmatrix} \Lambda_3 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$.

Proof of theorem 3 (only the first part is proved, the other proofs are similar)**.**

The optimal value of the GMM objective function reads

$$\begin{aligned}
G(\hat{\alpha},\hat{\beta}) &= vec(\sum_{t=1}^{T} \hat{\varepsilon}_t x_{t-1}')'((\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma^{-1}) vec(\sum_{t=1}^{T} \hat{\varepsilon}_t x_{t-1}') \\
&= vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1}(\sum_{t=1}^{T} \Delta x_t x_{t-1}')(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \\
&\quad \sum_{t=1}^{T} x_{t-1}x_{t-1}')'((\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma^{-1}) \\
&\quad vec(\sum_{t=1}^{T} (\Delta x_t - \hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1}(\sum_{t=1}^{T} \Delta x_t x_{t-1}')(\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \\
&\quad \sum_{t=1}^{T} x_{t-1}x_{t-1}') \\
&= vec((\Sigma^{-1} - \Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1})(\sum_{t=1}^{T} \Delta x_t x_{t-1}'))' \\
&\quad ((\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma) vec((\Sigma^{-1} - \Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1}) \\
&\quad (\sum_{t=1}^{T} \Delta x_t x_{t-1}'))
\end{aligned}$$

This functional consists of two parts, $((\sum_{t=1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma)$ and $vec((\Sigma^{-1} - \Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1})(\sum_{t=1}^{T} \Delta x_t x_{t-1}'))$, each of which limiting behavior is analyzed separately. Starting with the latter expression,

$$\frac{1}{T} vec((\Sigma^{-1} - \Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1})(\sum_{t=1}^{T} \Delta x_t x_{t-1}'))$$

$$
\begin{aligned}
&= \frac{1}{T} vec((\hat{\alpha}_\perp(\hat{\alpha}'_\perp \Sigma \hat{\alpha}_\perp)^{-1}\hat{\alpha}'_\perp(\sum_{t=1}^{T}\Delta x_t x'_{t-1})) \\
&\Rightarrow \frac{1}{T} vec(\alpha_\perp(\alpha'_\perp\Sigma\alpha_\perp)^{-1}\sum_{t=1}^{T}(\alpha'_\perp\Sigma^{\frac{1}{2}}\xi_t(\sum_{j=1}^{t-1}\xi'_j\Sigma^{\frac{1}{2}}\alpha_\perp(\beta'_\perp\alpha_\perp)^{-1}\beta'_\perp)) \\
&\Rightarrow (\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\otimes\alpha_\perp(\alpha'_\perp\Sigma\alpha_\perp)^{-1})vec(\Lambda_1(\int W_1 dW'_1)'\Lambda'_1)
\end{aligned}
$$

While

$$
\begin{aligned}
& T^2((\sum_{t=1}^{T}x_{t-1}x'_{t-1})^{-1}\otimes\Sigma) \\
\Rightarrow\ & (\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W'_1)^{-1}\Lambda_1^{-1}\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\otimes\Sigma).
\end{aligned}
$$

So

$$
\begin{aligned}
& G(\hat{\alpha},\hat{\beta}) \\
\Rightarrow\ & vec(\Lambda_1(\int W_1dW'_1)'\Lambda'_1)'(\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\otimes\alpha_\perp(\alpha'_\perp\Sigma\alpha_\perp)^{-1})' \\
& (\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int W_1W'_1)^{-1}\Lambda_1^{-1}\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\otimes\Sigma) \\
& (\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\otimes\alpha_\perp(\alpha'_\perp\Sigma\alpha_\perp)^{-1})vec(\Lambda_1(\int W_1dW'_1)'\Lambda'_1) \\
\Rightarrow\ & vec(\Lambda_1(\int W_1dW'_1)'\Lambda'_1)'(\Lambda_1^{-1\prime}(\int W_1W'_1)^{-1}\Lambda_1^{-1}\otimes(\alpha'_\perp\Sigma\alpha_\perp)^{-1}) \\
& vec(\Lambda_1(\int W_1dW'_1)'\Lambda'_1) \\
\Rightarrow\ & vec((\int W_1dW'_1)')'((\int W_1W'_1)^{-1}\otimes I_{k-r})vec((\int W_1dW'_1)') \\
\Rightarrow\ & tr[(\int W_1dW'_1)'(\int W_1W'_1)^{-1}(\int W_1dW'_1)]
\end{aligned}
$$

Proof of theorem 4.

The GMM objective function reads,

$$
\begin{aligned}
G(\alpha,\beta) &= vec(\sum_{t=1}^{T_1}\Sigma_1^{-1}\varepsilon_t x'_{t-1}+\sum_{t=T_1+1}^{T}\Sigma_2^{-1}\varepsilon_t x'_{t-1})' \\
&\quad (\sum_{t=1}^{T_1}(x_{t-1}x'_{t-1}\otimes\Sigma_1^{-1})+\sum_{t=T_1+1}^{T}(x_{t-1}x'_{t-1}\otimes\Sigma_2^{-1}))^{-1} \\
&\quad vec(\sum_{t=1}^{T_1}\Sigma_1^{-1}\varepsilon_t x'_{t-1}+\sum_{t=T_1+1}^{T}\Sigma_2^{-1}\varepsilon_t x'_{t-1})
\end{aligned}
$$

such that its derivative to $vec(\beta')$ becomes,

$$\begin{aligned}
& \frac{\partial G}{\partial vec(\beta')'} \\
= \quad & (I_k \otimes \alpha')((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1})) \\
& vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}\varepsilon_t x'_{t-1} + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\varepsilon_t x'_{t-1}) \\
= \quad & -((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_1^{-1}\alpha) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_2^{-1}\alpha))vec(\beta') \\
& vec(\sum_{t=1}^{T_1} \alpha'\Sigma_1^{-1}\Delta x_t x'_{t-1} + \sum_{t=T_1+1}^{T} \alpha'\Sigma_2^{-1}\Delta x_t x'_{t-1}) \\
\rightarrow \quad & vec(\beta') = ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_1^{-1}\alpha) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_2^{-1}\alpha))^{-1} \\
& vec(\sum_{t=1}^{T_1} \alpha'\Sigma_1^{-1}\Delta x_t x'_{t-1} + \sum_{t=T_1+1}^{T} \alpha'\Sigma_2^{-1}\Delta x_t x'_{t-1})
\end{aligned}$$

Estimators for $\hat{\alpha}$ and $\hat{\beta}$ then are

$$\begin{aligned}
vec(\hat{\alpha}) \quad = \quad & ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1}))^{-1} \\
& vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}\Delta x_t x'_{t-1} + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\Delta x_t x'_{t-1}) \\
= \quad & ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1}))^{-1} \\
& [(\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1})vec(\hat{a}_1) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1})vec(\hat{a}_2)]
\end{aligned}$$

and

$$\begin{aligned}
vec(\hat{\beta}') \quad = \quad & ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_1^{-1}\hat{\alpha}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_2^{-1}\hat{\alpha}))^{-1} \\
& vec(\sum_{t=1}^{T_1} \hat{\alpha}'\Sigma_1^{-1}\Delta x_t x'_{t-1} + \sum_{t=T_1+1}^{T} \hat{\alpha}'\Sigma_2^{-1}\Delta x_t x'_{t-1}) \\
= \quad & ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_1^{-1}\hat{\alpha}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_2^{-1}\hat{\alpha}))^{-1}
\end{aligned}$$

$$
\begin{aligned}
& [(\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_1^{-1}\hat{\alpha})vec(\hat{b}_1) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_2^{-1}\hat{\alpha})vec(\hat{b}_2)] \\
= \quad & vec(\begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix}') + \begin{pmatrix} 0 \\ ((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_1^{-1}\hat{\alpha}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \hat{\alpha}'\Sigma_2^{-1}\hat{\alpha}))_2^{-1} \end{pmatrix} \\
& vec(\sum_{t=1}^{T_1} \hat{\alpha}'\Sigma_1^{-1}\varepsilon_t x'_{t-1} + \sum_{t=T_1+1}^{T} \hat{\alpha}'\Sigma_2^{-1}\varepsilon_t x'_{t-1})
\end{aligned}
$$

where

$$
\hat{a}_1 = (\sum_{t=1}^{T_1} \Delta x_t x'_{t-1})(\sum_{t=1}^{T_1} x_{t-1}x'_{t-1})^{-1},
$$

$$
\hat{a}_2 = (\sum_{t=T_1+1}^{T} \Delta x_t x'_{t-1})(\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1})^{-1},
$$

$$
\begin{aligned}
\hat{b}_1 &= (\sum_{t=1}^{T_1} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{T_1} x_{t-1}\Delta x'_t)\Sigma_1^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma_1^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} \hat{a}_1 \\ (\sum_{t=1}^{T_1} x_{t-1}x'_{t-1})_2^{-1}(\sum_{t=1}^{T_1} x_{t-1}\Delta x'_t) \end{pmatrix} \Sigma_1^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma_1^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} \hat{b}_{11} \\ \hat{b}_{12} \end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\hat{b}_2 &= (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1})^{-1}(\sum_{t=T_1+1}^{T} x_{t-1}\Delta x'_t)\Sigma_2^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma_2^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} \hat{a}_2 \\ (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1})_2^{-1}(\sum_{t=T_1+1}^{T} x_{t-1}\Delta x'_t) \end{pmatrix} \Sigma_2^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma_2^{-1}\hat{\alpha})^{-1} \\
&= \begin{pmatrix} \hat{b}_{21} \\ \hat{b}_{22} \end{pmatrix},
\end{aligned}
$$

and $(\ )_2^{-1}$ represents the last $kr - r^2 = r(k-r)$ rows of $(\ )$.

Estimators for Σ_1 and Σ_2 can be obtained from the separate subsamples:

$$
\hat{\Sigma}_1 = \frac{1}{T_1 - k}\sum_{t=1}^{T_1} \Delta x_t(1 - x'_{t-1}(\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}x_{t-1})\Delta x'_t,
$$

$$
\hat{\Sigma}_2 = \frac{1}{T - T_1 - k}\sum_{t=T_1+1}^{T} \Delta x_t(1 - x'_{t-1}(\sum_{t=1}^{T} x_{t-1}x'_{t-1})^{-1}x_{t-1})\Delta x'_t.
$$

The limiting behavior of each of the four different parts of $vec(\hat{\beta}')$ will now be investigated. The crucial difference with the previous examples is the change of variance at point T_1, the stochastic trend of x_t, therefore, becomes

$\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'(\Sigma_1^{\frac{1}{2}}\sum_{j=1}^{\min(t,T_1)}\xi_j + I_k(t,T_1)\Sigma_2^{\frac{1}{2}}\sum_{j=1}^{t}\xi_j)$

where $I(t,T_1)$ is an indicator function, $I_k(t,T_1)=0,\ t\le T_1;\ I_k(t,T_1)=I_k,\ t>T_1$. If T_1 is such that $T_1=wT,\ (T-T_1)=(1-w)T,\ w=\frac{T_1}{T}$, the limiting behavior of x_t can be characterized by

$\frac{1}{\sqrt{T}}x_{Tl}=\frac{1}{\sqrt{T}}(\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'(\Sigma_1^{\frac{1}{2}}\sum_{j=1}^{\min(Tl,T_1)}\xi_j + I_k(Tl,T_1)\Sigma_2^{\frac{1}{2}}\sum_{j=T_1+1}^{Tl}\xi_j)\Rightarrow$

$\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}(\Lambda_1\int_0^{\min(w,l)}dW_1(t) + I_k(l,w)\Lambda_2\int_w^{l}dW_1(t))\Rightarrow$

$\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}(\Lambda_1 W_1(\min(w,l)) + I_k(l,w)\Lambda_2 W_1(l)),$

where $W_1(t),\ W_2(t)$ are stochastically independent $(k-r),\ r$ dimensional Brownian motions at time point $t,\ 0\le t\le 1;\ l,\ 0\le l\le 1$, and $\Lambda_1=(\alpha_\perp'\Sigma_1\alpha_\perp)^{\frac{1}{2}},\ \Lambda_2=(\alpha_\perp'\Sigma_2\alpha_\perp)^{\frac{1}{2}},\ \Omega_1=(\alpha'\Sigma_1\alpha)^{\frac{1}{2}},\ \Omega_2=(\alpha'\Sigma_2\alpha)^{\frac{1}{2}}$ (note that both Brownian motions appearing in the expression are independent).

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T_1}\hat{\alpha}'\Sigma_1^{-1}\varepsilon_t x_{t-1}'\\
=\ &\frac{1}{T}\sum_{t=1}^{T_1}\hat{\alpha}'\Sigma_1^{-\frac{1}{2}}\xi_t(\sum_{j=1}^{T_1}\xi_j)\Sigma_1^{\frac{1}{2}\prime}\alpha_\perp(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'\\
\Rightarrow\ &\Omega_1(\int_0^{w}dW_2W_1')\Lambda_1'(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{T}\sum_{t=T_1+1}^{T}\hat{\alpha}'\Sigma_2^{-1}\varepsilon_t x_{t-1}'\\
=\ &\frac{1}{T}\sum_{t=T_1+1}^{T}\hat{\alpha}'\Sigma_2^{-\frac{1}{2}}\xi_t(\sum_{j=1}^{T_1}\xi_j'\Sigma_1^{\frac{1}{2}\prime}+\sum_{j=T_1+1}^{t-1}\xi_j'\Sigma_2^{\frac{1}{2}\prime})\alpha_\perp(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'\\
\Rightarrow\ &\Omega_2(\int_w^{1}dW_2(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'
\end{aligned}$$

since both $\Sigma_2^{\frac{1}{2}}\alpha_\perp$ and $\Sigma_1^{\frac{1}{2}}\alpha_\perp\in span(\alpha_\perp)$ and $\Sigma_1^{-\frac{1}{2}}\alpha\in span(\alpha)$, $(\Sigma_2^{\frac{1}{2}}\alpha_\perp)'\Sigma_2^{-\frac{1}{2}}\alpha=0$ and $(\Sigma_1^{\frac{1}{2}}\alpha_\perp)'\Sigma_2^{-\frac{1}{2}}\alpha=0$. To obtain the limiting distribution of $\hat{\beta}$, we also need the following results,

$$T^2((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_1^{-1}\alpha) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \alpha'\Sigma_2^{-1}\alpha))_2^{-1}$$

$$\Rightarrow \quad ((\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \alpha'\Sigma_1^{-1}\alpha) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \alpha'\Sigma_2^{-1}\alpha))^{-1}$$

$$(\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp \otimes I_r)$$

So,

$$T[vec(\beta_2 - \hat{\beta}_2)]$$

$$\Rightarrow \quad ((\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \alpha'\Sigma_1^{-1}\alpha) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \alpha'\Sigma_2^{-1}\alpha))^{-1}$$

$$(\alpha'_\perp\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp \otimes I_r)(\beta_\perp(\alpha'_\perp\beta_\perp)^{-1} \otimes I_r)vec[\Omega_1(\int_0^w dW_2W'_1)\Lambda'_1 +$$

$$\Omega_2(\int_w^1 dW_2(t)(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt)]$$

$$\Rightarrow \quad ((\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \alpha'\Sigma_1^{-1}\alpha) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \alpha'\Sigma_2^{-1}\alpha))^{-1}$$

$$vec[\Omega_1(\int_0^w dW_2W'_1)\Lambda'_1 + \Omega_2(\int_w^1 dW_2(t)(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt)]$$

This limiting distribution is again normal as the Brownian motions in the stochastic integral W_1 and W_2 are stochastically independent. The limiting behavior of the optimal value of the GMM objective function can again be determined using the limiting behavior of the cointegrating vector estimator.

$$G(\hat{\alpha}, \hat{\beta}) \quad = \quad vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}\hat{\varepsilon}_t x'_{t-1} + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\hat{\varepsilon}_t x'_{t-1})'$$

$$
\begin{aligned}
& ((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))^{-1} \\
& vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}\hat{\varepsilon}_t x_{t-1}' + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\hat{\varepsilon}_t x_{t-1}') \\
= \quad & vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}' + \\
& \sum_{t=T_1+1}^{T} \Sigma_2^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}')'((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) \\
& +(\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))^{-1} vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}' \\
& + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}')
\end{aligned}
$$

Elements of this objective function are a.o.,

$$
\begin{aligned}
& ((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))^{-1} \\
& vec(\sum_{t=1}^{T_1} \Sigma_1^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}' + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}(\Delta x_t - \hat{\alpha}\hat{\beta}' x_{t-1})x_{t-1}') \\
= \quad & [((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))^{-1} - (I_k \otimes \hat{\alpha}) \\
& ((I_k \otimes \hat{\alpha})'((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))^{-1}(I_k \otimes \hat{\alpha}))^{-1} \\
& (I_k \otimes \hat{\alpha})'] vec((\sum_{t=1}^{T_1} \Sigma_1^{-1}\varepsilon_t x_{t-1}' + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\varepsilon_t x_{t-1}') \\
= \quad & (I_k \otimes \hat{\alpha}_\perp)((I_k \otimes \hat{\alpha}_\perp')((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1})) \\
& (I_k \otimes \hat{\alpha}_\perp))^{-1}(I_k \otimes \hat{\alpha}_\perp')vec((\sum_{t=1}^{T_1} \Sigma_1^{-1}\varepsilon_t x_{t-1}' + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\varepsilon_t x_{t-1}'),
\end{aligned}
$$

which has a limiting behavior following from,

$$
T(I_k \otimes \hat{\alpha}_\perp)((I_k \otimes \hat{\alpha}_\perp')((\sum_{t=1}^{T_1} x_{t-1}x_{t-1}' \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}' \otimes \Sigma_2^{-1}))
$$

$$(I_k \otimes \hat{\alpha}_\perp))^{-1}(I_k \otimes \hat{\alpha}'_\perp)vec((\sum_{t=1}^{T_1} \Sigma_1^{-\frac{1}{2}}\xi_t x'_{t-1} + \sum_{t=T_1+1}^{T} \Sigma_2^{-1}\varepsilon_t x'_{t-1})$$

$$\Rightarrow \quad T(I_k \otimes \alpha_\perp)((I_k \otimes \alpha'_\perp)((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1}))$$

$$(I_k \otimes \alpha_\perp))^{-1}vec([\sum_{t=1}^{T_1} \alpha'_\perp \Sigma_1^{-\frac{1}{2}}\xi_t(\sum_{j=1}^{t-1} \xi'_j \Sigma_1^{\frac{1}{2}\prime} + (\sum_{t=T_1+1}^{T} \alpha'_\perp \Sigma_2^{-\frac{1}{2}}\xi_t)$$

$$(\sum_{j=1}^{T_1} \xi'_j \Sigma_1^{\frac{1}{2}\prime} + \sum_{j=T_1+1}^{t-1} \xi'_j \Sigma_2^{\frac{1}{2}\prime})]\alpha_\perp(\beta'_\perp\alpha_\perp)^{-1}\beta'_\perp)$$

$$\Rightarrow \quad (\beta_\perp(\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp \otimes \alpha_\perp)((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \alpha'_\perp\Sigma_1^{-1}\alpha_\perp) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \alpha'_\perp\Sigma_2^{-1}\alpha_\perp))^{-1}$$

$$vec[\Lambda_1(\int_0^w dW_2W'_1)\Lambda'_1 + \Lambda_2(\int_w^1 dW_2(t)(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt)]$$

For determining the limiting behavior of the objective function we also need,

$$\frac{1}{T^2}((\sum_{t=1}^{T_1} x_{t-1}x'_{t-1} \otimes \Sigma_1^{-1}) + (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1} \otimes \Sigma_2^{-1}))$$

$$\Rightarrow \quad (\beta_\perp(\alpha'_\perp\beta_\perp)^{-1} \otimes I_k)((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \Sigma_1^{-1}) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \Sigma_2^{-1}))$$

$$((\beta'_\perp\alpha_\perp)^{-1}\beta'_\perp \otimes I_k).$$

So,

$$G(\hat{\alpha}, \hat{\beta})$$

$$\Rightarrow \quad vec[\Lambda_1(\int_0^w dW_2W'_1)\Lambda'_1 + \Lambda_2(\int_w^1 dW_2(t)(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt)]'$$

$$((\Lambda_1 \int_0^w W_1W'_1\Lambda'_1 \otimes \alpha'_\perp\Sigma_1^{-1}\alpha_\perp) +$$

$$(\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))(\Lambda_2 W_2(t) + \Lambda_1 W_1(w))' dt \otimes \alpha_\perp' \Sigma_2^{-1} \alpha_\perp))^{-1}$$
$$vec[\Lambda_1(\int_0^w dW_2 W_1')\Lambda_1' + \Lambda_2(\int_w^1 dW_2(t)(\Lambda_2 W_1(t) + \Lambda_1 W_1(w))' dt)]$$

$$\Rightarrow \quad vec[(\int_0^w dW_2 W_1') + \Lambda_1^{-1}\Lambda_2(\int_w^1 dW_2(t)(\Lambda_1^{-1}\Lambda_2 W_1(t) + W_1(w))' dt)]'$$
$$((\int_0^w W_1 W_1' \otimes I_{k-r}) + (\int_w^1 (\Lambda_1^{-1}\Lambda_2 W_1(t) + W_1(w))(\Lambda_1^{-1}\Lambda_2 W_2(t) +$$
$$W_1(w))' dt \otimes \Lambda_1^{-1\prime} \alpha_\perp' \Sigma_2^{-1} \alpha_\perp \Lambda_1^{-1}))^{-1}$$
$$vec[(\int_0^w dW_2 W_1') + \Lambda_1^{-1}\Lambda_2(\int_w^1 dW_2(t)(\Lambda_1^{-1}\Lambda_2 W_1(t) + W_1(w))' dt)]$$

Proof of theorem 5.

For the case of general kind of heteroscedasticity, the White covariance matrix estimator can be used in the GMM objective function. We analyze the behavior of the resulting estimator using this expression of the covariance matrix for the case analyzed previously, i.e. a change of variance at point T_1. The GMM objective function now becomes

$$\begin{aligned} G(\alpha, \beta) &= vec(\sum_{t=1}^T \varepsilon_t x_{t-1}')'(\sum_{t=1}^T (x_{t-1} x_{t-1}' \otimes \hat{\varepsilon}_t \hat{\varepsilon}_t'))^{-1} \\ & \quad vec(\sum_{t=1}^T \varepsilon_t x_{t-1}') \end{aligned}$$

The estimators of α and β then result from,

$$\begin{aligned} & \frac{\partial G}{\partial vec(\beta')'} \\ = \; & (\sum_{t=1}^T x_{t-1} x_{t-1}' \otimes I_k)(I_k \otimes \alpha')(\sum_{t=1}^T (x_{t-1} x_{t-1}' \otimes \hat{\varepsilon}_t \hat{\varepsilon}_t'))^{-1} \\ & vec(\sum_{t=1}^T \varepsilon_t x_{t-1}') \\ = \; & (\sum_{t=1}^T x_{t-1} x_{t-1}' \otimes I_k)(I_k \otimes \alpha')(\sum_{t=1}^T (x_{t-1} x_{t-1}' \otimes \hat{\varepsilon}_t \hat{\varepsilon}_t'))^{-1}[(I_k \otimes \alpha) \end{aligned}$$

$$(\sum_{t=1}^{T} x_{t-1}x'_{t-1} \otimes I_k)vec(\beta') - vec(\sum_{t=1}^{T} \Delta x_t x'_{t-1})]$$

So,

$$\begin{aligned}
vec(\hat{\beta}'_2) &= -(\sum_{t=1}^{T} x_{t-1}x'_{t-1} \otimes I_r)_2^{-1} \\
&\quad ((I_k \otimes \hat{\alpha}')(\sum_{t=1}^{T}(x_{t-1}x'_{t-1} \otimes \hat{\varepsilon}_t\hat{\varepsilon}'_t))^{-1}(I_k \otimes \hat{\alpha}))^{-1} \\
&\quad (I_k \otimes \hat{\alpha})(\sum_{t=1}^{T}(x_{t-1}x'_{t-1} \otimes \hat{\varepsilon}_t\hat{\varepsilon}'_t))^{-1}vec(\sum_{t=1}^{T} \Delta x_t x'_{t-1}) \\
&= vec(\beta'_2) - (\sum_{t=1}^{T} x_{t-1}x'_{t-1} \otimes I_r)_2^{-1}((I_k \otimes \hat{\alpha}')(\sum_{t=1}^{T}(x_{t-1}x'_{t-1} \otimes \hat{\varepsilon}_t\hat{\varepsilon}'_t))^{-1}(I_k \otimes \hat{\alpha}))^{-1} \\
&\quad (I_k \otimes \hat{\alpha})(\sum_{t=1}^{T}(x_{t-1}x'_{t-1} \otimes \hat{\varepsilon}_t\hat{\varepsilon}'_t))^{-1}vec(\sum_{t=1}^{T} \varepsilon_t x'_{t-1})
\end{aligned}$$

and

$$\begin{aligned}
\hat{\alpha} &= (\sum_{t=1}^{T} \Delta x_t(1 - x_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1}) \\
&\quad (\sum_{t=1}^{T} x_{1t-1}(1 - x_{2t-1}(\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1})^{-1}.
\end{aligned}$$

As we assume a change of variance at time T_1, the stochastic trend in x_t is identical to the one in the previous case,

$\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp(\Sigma_1^{\frac{1}{2}} \sum_{j=1}^{\min(t,T_1)} \xi_j + I_k(t, T_1)\Sigma_2^{\frac{1}{2}} \sum_{j=T_1+1}^{t} \xi_j)$

where the definitions used in the proof of theorem 4 are used.

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t x'_{t-1} \\
= \;& \frac{1}{T}(\sum_{t=1}^{T_1} \Sigma_1^{\frac{1}{2}}\xi_t(\sum_{j=1}^{t-1} \xi'_j\Sigma_1^{\frac{1}{2}\prime}) + \sum_{t=T_1+1}^{T} \Sigma_2^{\frac{1}{2}}\xi_t(\sum_{j=1}^{T_1} \xi'_j\Sigma_1^{\frac{1}{2}\prime} + \sum_{j=T_1+1}^{t-1} \xi'_j\Sigma_1^{\frac{1}{2}\prime})) \\
& \alpha_\perp(\beta'_\perp\alpha_\perp)^{-1}\beta'_\perp \\
\Rightarrow \;& [\Sigma_1^{\frac{1}{2}} \int_0^w dW_2(t)(\Lambda_1 W_1(t))'dt + \Sigma_2^{\frac{1}{2}} \int_w^1 dW_2(t)(\Lambda_1 W_1(w) + \Lambda_2 W_1(t))'dt] \\
& (\beta'_\perp\alpha_\perp)^{-1}\beta'_\perp
\end{aligned}$$

Using lemma 1, it follows that

$$\begin{aligned}
&\frac{1}{T^2}(\sum_{t=1}^{T}(x_{t-1}x_{t-1}'\otimes\hat{\varepsilon}_t\hat{\varepsilon}_t'))\\
\Rightarrow\quad &\frac{1}{T^2}((\sum_{t=1}^{T_1}x_{t-1}x_{t-1}'\otimes\Sigma_1)+(\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}'\otimes\Sigma_2))\\
\Rightarrow\quad &(\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\otimes I_k)((\Lambda_1\int_0^w W_1W_1'\Lambda_1'\otimes\Sigma_1)+\\
&(\int_w^1(\Lambda_2W_1(t)+\Lambda_1W_1(w))(\Lambda_2W_2(t)+\Lambda_1W_1(w))'dt\otimes\Sigma_2))\\
&((\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'\otimes I_k)
\end{aligned}$$

Such that the inverse has a limiting behavior characterized by,

$$\begin{aligned}
&T^2(\sum_{t=1}^{T}(x_{t-1}x_{t-1}'\otimes\hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}\\
\Rightarrow\quad &(\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\otimes I_k)((\Lambda_1\int_0^w W_1W_1'\Lambda_1'\otimes\Sigma_1)+\\
&(\int_w^1(\Lambda_2W_1(t)+\Lambda_1W_1(w))(\Lambda_2W_2(t)+\Lambda_1W_1(w))'dt\otimes\Sigma_2))^{-1}\\
&(\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\otimes I_k).
\end{aligned}$$

So,

$$\begin{aligned}
&T(I_k\otimes\hat{\alpha}')(\sum_{t=1}^{T}(x_{t-1}x_{t-1}'\otimes\hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}vec(\sum_{t=1}^{T}\varepsilon_tx_{t-1}')\\
\Rightarrow\quad &(I_k\otimes\alpha')(\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\otimes I_k)((\Lambda_1\int_0^w W_1W_1'\Lambda_1'\otimes\Sigma_1)+\\
&(\int_w^1(\Lambda_2W_1(t)+\Lambda_1W_1(w))(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt\otimes\Sigma_2))^{-1}\\
&(\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\otimes I_k)(\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\otimes I_k)\\
&vec[\Sigma_1^{\frac{1}{2}}\int_0^w dW_2(t)(\Lambda_1W_1(t))'dt+\Sigma_2^{\frac{1}{2}}\int_w^1 dW_2(t)(\Lambda_1W_1(w)+\Lambda_2W_1(t))'dt]
\end{aligned}$$

$$\Rightarrow \quad (\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes \alpha')((\Lambda_1 \int_0^w W_1W_1'\Lambda_1' \otimes \Sigma_1) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt \otimes \Sigma_2))^{-1}$$

$$vec[\Sigma_1^{\frac{1}{2}} \int_0^w dW_2(t)(\Lambda_1W_1(t))'dt + \Sigma_2^{\frac{1}{2}} \int_w^1 dW_2(t)(\Lambda_1W_1(w) + \Lambda_2W_1(t))'dt]$$

$$\Rightarrow \quad n(0, (\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes \alpha')((\Lambda_1 \int_0^w W_1W_1'\Lambda_1' \otimes \Sigma_1) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_1(t) + \Lambda_1W_1(w))'dt \otimes \Sigma_2))^{-1}$$

$$(\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' \otimes \alpha))$$

Also the limiting behavior of some other matrices is needed,

$$T^2(I_k \otimes \hat{\alpha}')(\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}(I_k \otimes \hat{\alpha})$$

$$\Rightarrow \quad (\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes \alpha')((\Lambda_1 \int_0^w W_1W_1'\Lambda_1' \otimes \Sigma_1) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes \Sigma_2))^{-1}$$

$$(\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' \otimes \alpha),$$

and

$$T^2(\sum_{t=1}^T x_{t-1}x_{t-1}' \otimes I_r)_2^{-1}$$

$$\Rightarrow \quad ((\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1W_1'\Lambda_1' \otimes I_r) +$$

$$(\int_w^1 (\Lambda_2W_1(t) + \Lambda_1W_1(w))(\Lambda_2W_2(t) + \Lambda_1W_1(w))'dt \otimes I_r))^{-1}$$

$$(\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' \otimes I_r).$$

So,

$$
\begin{aligned}
& T(vec(\beta_2' - \hat{\beta}_2') \\
\Rightarrow\ & ((\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp \otimes I_r)((\Lambda_1 \int_0^w W_1 W_1'\Lambda_1' \otimes I_r) + \\
& (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))(\Lambda_2 W_2(t) + \Lambda_1 W_1(w))' dt \otimes I_r))^{-1} \\
& [(I_{k-r} \otimes \alpha')((\Lambda_1 \int_0^w W_1 W_1'\Lambda_1' \otimes \Sigma_1) + \\
& (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))(\Lambda_2 W_2(t) + \Lambda_1 W_1(w))' dt \otimes \Sigma_2))^{-1} \\
& (I_{k-r} \otimes \alpha)]^{-1}((\Lambda_1 \int_0^w W_1 W_1'\Lambda_1' \otimes I_r) + \\
& (\int_w^1 (\Lambda_2 W_1(t) + \Lambda_1 W_1(w))(\Lambda_2 W_2(t) + \Lambda_1 W_1(w))' dt \otimes I_r))^{-1} \\
& (\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1} \otimes I_r)
\end{aligned}
$$

The optimal value of the objective function reads,

$$
\begin{aligned}
G(\hat{\alpha}, \hat{\beta}) &= vec(\sum_{t=1}^T \hat{\varepsilon}_t x_{t-1}')'(\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1} \\
& \quad vec(\sum_{t=1}^T \hat{\varepsilon}_t x_{t-1}'),
\end{aligned}
$$

the limiting behavior of its different elements is now determined,

$$
\begin{aligned}
& (\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1} vec(\sum_{t=1}^T \hat{\varepsilon}_t x_{t-1}') \\
= \ & [(\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1} - (\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}(I_k \otimes \hat{\alpha}) \\
& ((I_k \otimes \hat{\alpha}')(\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1}(I_k \otimes \hat{\alpha}))^{-1} \\
& (I_k \otimes \hat{\alpha})(\sum_{t=1}^T (x_{t-1}x_{t-1}' \otimes \hat{\varepsilon}_t\hat{\varepsilon}_t'))^{-1} vec(\sum_{t=1}^T \Delta x_t x_{t-1}')
\end{aligned}
$$

$$= (I_k \otimes \hat{\alpha}_\perp)((I_k \otimes \hat{\alpha}'_\perp)(\sum_{t=1}^{T}(x_{t-1}x'_{t-1} \otimes \hat{\varepsilon}_t\hat{\varepsilon}'_t))(I_k \otimes \hat{\alpha}_\perp))^{-1}$$
$$(I_k \otimes \hat{\alpha}'_\perp)vec(\sum_{t=1}^{T}\varepsilon_t x'_{t-1})$$

and

$$\frac{1}{T}(I_k \otimes \hat{\alpha}'_\perp)vec(\sum_{t=1}^{T}\varepsilon_t x'_{t-1})$$
$$\Rightarrow \frac{1}{T}(vec(\sum_{t=1}^{T}\alpha'_\perp\varepsilon_t x'_{t-1})$$
$$\Rightarrow vec[\Lambda_1\int_{\mathbf{0}}^{w}dW_1(t)(\Lambda_1W_1(t))'dt + \Lambda_2\int_{w}^{1}dW_1(t)(\Lambda_1W_1(w)+\Lambda_2W_1(t))'dt].$$

So that,

$$G(\hat{\alpha},\hat{\beta})$$
$$\Rightarrow vec[\Lambda_1(\int_{\mathbf{0}}^{w}dW_1W'_1)\Lambda'_1 + \Lambda_2(\int_{w}^{1}dW_1(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)]'$$
$$((\Lambda_1\int_{\mathbf{0}}^{w}W_1W'_1\Lambda'_1 \otimes \alpha'_\perp\Sigma_1^{-1}\alpha_\perp) +$$
$$(\int_{w}^{1}(\Lambda_2W_1(t)+\Lambda_1W_1(w))(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt \otimes \alpha'_\perp\Sigma_2^{-1}\alpha_\perp))^{-1}$$
$$vec[\Lambda_1(\int_{\mathbf{0}}^{w}dW_1W'_1)\Lambda'_1 + \Lambda_2(\int_{w}^{1}dW_1(t)(\Lambda_2W_1(t)+\Lambda_1W_1(w))'dt)]$$

$$\Rightarrow vec[(\int_{\mathbf{0}}^{w}dW_1W'_1) + \Lambda_1^{-1}\Lambda_2(\int_{w}^{1}dW_1(t)(\Lambda_1^{-1}\Lambda_2W_1(t)+W_1(w))'dt)]'$$
$$((\int_{\mathbf{0}}^{w}W_1W'_1 \otimes I_{k-r}) + (\int_{w}^{1}(\Lambda_1^{-1}\Lambda_2W_1(t)+W_1(w))(\Lambda_1^{-1}\Lambda_2W_2(t)+$$
$$W_1(w))'dt \otimes \Lambda_1^{-1\prime}\alpha'_\perp\Sigma_2^{-1}\alpha_\perp\Lambda_1^{-1}))^{-1}$$
$$vec[(\int_{\mathbf{0}}^{w}dW_1W'_1) + \Lambda_1^{-1}\Lambda_2(\int_{w}^{1}dW_1(t)(\Lambda_1^{-1}\Lambda_2W_1(t)+W_1(w))'dt)].$$

Proof of theorem 6.

In this part, we investigate the influence of a change in the value of the multiplicator, α, and cointegrating vector, β, at point T_1. The model therefore is

$$\begin{aligned} \Delta x_t &= \alpha\beta' x_{t-1} + \varepsilon_t && t = 1, ..., T_1; \\ \Delta x_t &= \theta\gamma' x_{t-1} + \varepsilon_t && t = T_1 + 1, ..., T. \end{aligned}$$

where $\gamma = \beta$. We now derive the stochastic trend specification of the second part dataset generated by the model,

$$\begin{aligned} \Delta x_t &= \theta\gamma' x_{t-1} + \varepsilon_t \quad \Leftrightarrow \\ \begin{pmatrix} \gamma' x_t \\ \Delta\gamma'_\perp x_t \end{pmatrix} &= \begin{pmatrix} I_r + \gamma'\theta \\ \gamma'_\perp\theta \end{pmatrix} \gamma' x_{t-1} + \begin{pmatrix} \gamma' \\ \gamma'_\perp \end{pmatrix} \varepsilon_t \Leftrightarrow \\ \begin{pmatrix} \gamma' x_t \\ \Delta\gamma'_\perp x_t \end{pmatrix} &= \begin{pmatrix} \gamma' \\ \gamma'_\perp \end{pmatrix} \varepsilon_t + \begin{pmatrix} I_r + \gamma'\theta \\ \gamma'_\perp\theta \end{pmatrix} \sum_{i=1}^{t-T_1-1} (I_r + \gamma'\theta)^{i-1}\gamma'\varepsilon_{t-i} \\ &\quad + \begin{pmatrix} I_r + \gamma'\theta \\ \gamma'_\perp\theta \end{pmatrix} (I_r + \gamma'\theta)^{t-T_1-1}\gamma' x_{T_1} \\ &= \begin{pmatrix} \gamma' \\ \gamma'_\perp \end{pmatrix} \varepsilon_t + \begin{pmatrix} I_r + \gamma'\theta \\ \gamma'_\perp\theta \end{pmatrix} \sum_{i=1}^{t-T_1-1} (I_r + \gamma'\theta)^{i-1}\gamma'\varepsilon_{t-i} \\ &\quad + \begin{pmatrix} I_r + \gamma'\theta \\ \gamma'_\perp\theta \end{pmatrix} (I_r + \gamma'\theta)^{t-T_1-1} \sum_{i=t-T_1}^{\infty} (I_r + \beta'\alpha)^{i-1}\gamma'\varepsilon_{t-i} \end{aligned}$$

Since $(I_r + \gamma'\theta)^{t-T_1}$ converges to 0 when $(t - T_1) \to \infty$, we neglect the stochastic trend resulting from the latter part (we also rid of the stochastic trend in the first differences in this way) and let the stochastic trend result from $\gamma'_\perp x_{T_1+1} = \gamma'_\perp x_{T_1} + \gamma'_\perp \varepsilon_{T_1+1}$.

$$\begin{aligned} \frac{1}{\sqrt{T}} x_{Tl} &= \frac{1}{\sqrt{T}} (\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp \Sigma^{\frac{1}{2}} \sum_{j=1}^{T} \xi_j && Tl \le T_1 \\ &= \frac{1}{\sqrt{T}} (\gamma_\perp(\gamma'_\perp\gamma_\perp)^{-1}\gamma'_\perp\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp \Sigma^{\frac{1}{2}} \sum_{j=1}^{T_1} \xi_j && Tl > T_1 \\ &\quad + \gamma_\perp(\theta'_\perp\gamma_\perp)^{-1}\theta'_\perp \Sigma^{\frac{1}{2}} \sum_{j=T_1+1}^{Tl} \xi_j) \end{aligned}$$

If T_1 is such that $T_1 = wT$, $(T - T_1) = (1-w)T$, $w = \frac{T_1}{T}$, $0 \leq w \leq 1$; $0 \leq l \leq 1$, the limiting behavior of x_t can be characterized by

$$\begin{aligned}
\frac{1}{\sqrt{T}} x_{Tl} &= \frac{1}{\sqrt{T}} (\beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \alpha_\perp' \Sigma^{\frac{1}{2}} \sum_{j=1}^{Tl} \xi_j && Tl \leq T_1 \\
&\Rightarrow \beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \Lambda_1 \int_0^l dW_1(t) \\
&= \frac{1}{\sqrt{T}} (\gamma_\perp (\gamma_\perp' \gamma_\perp)^{-1} \gamma_\perp' \beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \alpha_\perp' \Sigma^{\frac{1}{2}} \sum_{j=1}^{T_1} \xi_j && Tl > T_1 \\
&\quad + \gamma_\perp (\theta_\perp' \gamma_\perp)^{-1} \theta_\perp' \Sigma^{\frac{1}{2}} \sum_{j=T_1+1}^{Tl} \xi_j) \\
&\Rightarrow \gamma_\perp (\gamma_\perp' \gamma_\perp)^{-1} \gamma_\perp' \beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \Lambda_1 \int_0^w dW_1(t) \\
&\quad + \gamma_\perp (\theta_\perp' \gamma_\perp)^{-1} \Lambda_2 \int_w^l dW_1(t)) \\
&\Rightarrow \gamma_\perp (\gamma_\perp' \gamma_\perp)^{-1} \gamma_\perp' [\beta_\perp (\alpha_\perp' \beta_\perp)^{-1} \Lambda_1 W_1(w) \\
&\quad + \gamma_\perp (\theta_\perp' \gamma_\perp)^{-1} \Lambda_2 W_1(l-w)]
\end{aligned}$$

where $W_1(t)$ is a $(k-r)$ dimensional Brownian motions at time point t, $0 \leq t \leq 1$; l, $0 \leq l \leq 1$, and $\Lambda_1 = (\alpha_\perp' \Sigma \alpha_\perp)^{\frac{1}{2}}$, $\Lambda_2 = (\theta_\perp' \Sigma \theta_\perp)^{\frac{1}{2}}$, $\Omega_1 = (\alpha' \Sigma^{-1} \alpha)^{\frac{1}{2}}$, $\Omega_2 = (\theta' \Sigma^{-1} \theta)^{\frac{1}{2}}$. The GMM objective function does in this case read

$$\begin{aligned}
G(\alpha, \beta, \gamma, \theta) &= vec(\sum_{t=1}^{T_1} \varepsilon_t x_{t-1}', \sum_{t=T_1+1}^{T} \varepsilon_t x_{t-1}')' \\
&\quad \begin{pmatrix} ((\sum_{t=1}^{T_1} x_{t-1} x_{t-1}')^{-1} \otimes \Sigma^{-1}) & 0 \\ 0 & ((\sum_{t=T_1+1}^{T} x_{t-1} x_{t-1}')^{-1} \otimes \Sigma^{-1}) \end{pmatrix} \\
&\quad vec(\sum_{t=1}^{T_1} \varepsilon_t x_{t-1}', \sum_{t=T_1+1}^{T} \varepsilon_t x_{t-1}')
\end{aligned}$$

where $vec(A, B) = (vec(A)'\ vec(B)')'$. The cointegrating vector estimators, $\hat{\beta}$ and $\hat{\gamma}$, and multiplicators estimators , $\hat{\alpha}$ and $\hat{\theta}$, are all identical to these estimators in the standard case,

$$\hat{\alpha} = (\sum_{t=1}^{T_1} \Delta x_t (1 - x_{2t-1}' (\sum_{t=1}^{T_1} x_{2t-1} x_{2t-1}')^{-1} x_{2t-1}) x_{1t-1}')$$

$$
\begin{aligned}
& (\sum_{t=1}^{T_1} x_{1t-1}(1 - x'_{2t-1}(\sum_{t=1}^{T_1} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1})^{-1} \\
= \; & (\sum_{t=1}^{T_1} \Delta x_t(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})')(\sum_{t=1}^{T_1}(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})(x_{1t-1} - \tilde{\beta}'_2 x_{2t-1})')^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\hat{\theta} \; = \; & (\sum_{t=T_1+1}^{T} \Delta x_t(1 - x'_{2t-1}(\sum_{t=T_1+1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1}) \\
& (\sum_{t=T_1+1}^{T} x_{1t-1}(1 - x'_{2t-1}(\sum_{t=T_1+1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1})x'_{1t-1})^{-1} \\
= \; & (\sum_{t=T_1+1}^{T} \Delta x_t(x_{1t-1} - \tilde{\gamma}'_2 x_{2t-1})')(\sum_{t=T_1+1}^{T}(x_{1t-1} - \tilde{\gamma}'_2 x_{2t-1})(x_{1t-1} - \tilde{\gamma}'_2 x_{2t-1})')^{-1},
\end{aligned}
$$

with $\tilde{\beta}_2 = (\sum_{t=1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1}x'_{1t-1}$, $\tilde{\gamma}_2 = (\sum_{t=T_1+1}^{T} x_{2t-1}x'_{2t-1})^{-1}x_{2t-1}x'_{1t-1}$.
The limiting distributions of $\hat{\alpha}$ and $\hat{\theta}$ read,

$$
\begin{aligned}
\sqrt{T}(\hat{\alpha} - \alpha) \; & \Rightarrow \; n(0,\ cov(\beta' x)^{-1} \otimes w\Sigma), \\
\sqrt{T}(\hat{\theta} - \theta) \; & \Rightarrow \; n(0,\ cov(\gamma' x)^{-1} \otimes (1-w)\Sigma).
\end{aligned}
$$

With respect to the cointegrating vector,

$$
\begin{aligned}
\hat{\beta} \; & = \; (\sum_{t=1}^{T_1} x_{t-1}x'_{t-1})^{-1}(\sum_{t=1}^{T_1} x_{t-1}\Delta x'_t)\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1} \\
\hat{\gamma} \; & = \; (\sum_{t=T_1+1}^{T} x_{t-1}x'_{t-1})^{-1}(\sum_{t=T_1+1}^{T} x_{t-1}\Delta x'_t)\Sigma^{-1}\hat{\theta}(\hat{\theta}'\Sigma^{-1}\hat{\theta})^{-1}
\end{aligned}
$$

The limiting distribtutions of these cointegrating vector estimators are,

$$
\begin{aligned}
T(\hat{\beta} - \beta) \; & \Rightarrow \; \begin{pmatrix} 0 \\ (\beta'_\perp\beta_\perp)^{-1}\beta'_\perp\alpha_\perp\Lambda_1^{-1\prime}(\int_0^w W_1W'_1)^{-1}(\int_0^w W_1 dW'_2)\Omega'_1 \end{pmatrix} \\
T(\gamma_2 - \hat{\gamma}_2) \; & \Rightarrow \; [(\int_w^1 (\gamma'_\perp\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma'_\perp\gamma_\perp(\theta'_\perp\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& (\gamma'_\perp\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma'_\perp\gamma_\perp(\theta'_\perp\gamma_\perp)^{-1}\Lambda_2 W_1(t))'dt)]^{-1} \\
& \int_w^1 [\gamma'_\perp\beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma'_\perp\gamma_\perp(\theta'_\perp\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& dW_2(t)'dt]\Omega'_2
\end{aligned}
$$

The optimal value of the GMM objective function becomes,

$$\begin{aligned} G(\hat{\alpha},\hat{\beta},\hat{\gamma},\hat{\theta}) &= vec(\sum_{t=1}^{T_1}\hat{\varepsilon}_t x_{t-1}')'((\sum_{t=1}^{T_1}x_{t-1}x_{t-1}')^{-1}\otimes\Sigma^{-1})vec(\sum_{t=1}^{T_1}\hat{\varepsilon}_t x_{t-1}') \\ &\quad +vec(\sum_{t=T_1+1}^{T}\hat{\varepsilon}_t x_{t-1}')'((\sum_{t=T_1+1}^{T}x_{t-1}x_{t-1}')^{-1}\otimes\Sigma^{-1}) \\ &\quad vec(\sum_{t=T_1+1}^{T}\hat{\varepsilon}_t x_{t-1}'), \end{aligned}$$

whose elements show the limiting behavior,

$$\begin{aligned} &\frac{1}{T}vec(\sum_{t=1}^{T_1}\Sigma^{-1}\hat{\varepsilon}_t x_{t-1}') \\ = \;& \frac{1}{T}vec((\Sigma^{-1}-\Sigma^{-1}\hat{\alpha}(\hat{\alpha}'\Sigma^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\Sigma^{-1})(\sum_{t=1}^{T}\Delta x_t x_{t-1}')) \\ \Rightarrow \;& \frac{1}{T}vec(\alpha_\perp(\alpha_\perp'\Sigma\alpha_\perp)^{-1}\alpha_\perp'\Lambda_1(\int_0^w W_1 dW_1')'\Lambda_1')(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp'), \end{aligned}$$

The first component in the sum contained in the objective function shows the limiting behavior,

$$\begin{aligned} &vec(\sum_{t=1}^{T_1}\Sigma^{-1}\hat{\varepsilon}_t x_{t-1}')'((\sum_{t=1}^{T_1}x_{t-1}x_{t-1}')^{-1}\otimes\Sigma)vec(\sum_{t=1}^{T_1}\Sigma^{-1}\hat{\varepsilon}_t x_{t-1}') \\ \Rightarrow \;& vec(\alpha_\perp(\alpha_\perp'\Sigma\alpha_\perp)^{-1}\alpha_\perp'\Lambda_1(\int_0^w W_1 dW_1')'\Lambda_1')(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp')' \\ & (\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\alpha_\perp\Lambda_1^{-1}(\int_0^w W_1 W_1')^{-1}\Lambda_1^{-1\prime}\alpha_\perp'\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\otimes\Sigma) \\ & vec(\alpha_\perp(\alpha_\perp'\Sigma\alpha_\perp)^{-1}\alpha_\perp'\Lambda_1(\int_0^w W_1 dW_1')'\Lambda_1')(\beta_\perp'\alpha_\perp)^{-1}\beta_\perp') \\ \Rightarrow \;& vec((\int_0^w W_1 dW_1')')'((\int_0^w W_1 W_1')^{-1}\otimes I_{k-r})vec((\int_0^w W_1 dW_1')'), \end{aligned}$$

and for the second sum we need,

$$\frac{1}{T}vec(\sum_{t=T_1+1}^{T}\Sigma^{-1}\hat{\varepsilon}_t x_{t-1}')$$

$$
\begin{aligned}
&= \frac{1}{T} vec((\Sigma^{-1} - \Sigma^{-1}\hat{\theta}(\hat{\theta}'\Sigma^{-1}\hat{\theta})^{-1}\hat{\theta}'\Sigma^{-1})(\sum_{t=1}^{T} \Delta x_t x_{t-1}')) \\
&\Rightarrow \; vec(\theta_\perp(\theta_\perp'\Sigma\theta_\perp)^{-1}\theta_\perp'\Lambda_2 \int_w^1 [\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) \\
&\quad + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')'(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'),
\end{aligned}
$$

such that it shows the limiting behavior,

$$
\begin{aligned}
& vec(\sum_{t=T_1+1}^{T} \Sigma^{-1}\hat{\varepsilon}_t x_{t-1}')'((\sum_{t=T_1+1}^{T} x_{t-1}x_{t-1}')^{-1} \otimes \Sigma) vec(\sum_{t=T_1+1}^{T} \Sigma^{-1}\hat{\varepsilon}_t x_{t-1}') \\
\Rightarrow \; & vec(\theta_\perp(\theta_\perp'\Sigma\theta_\perp)^{-1}\theta_\perp'\Lambda_2 \int_w^1 [\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \\
& \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')'(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp')' \\
& (\gamma_\perp[\int_w^1 (\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& (\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))'dt]^{-1}\gamma_\perp' \\
& \otimes\Sigma) vec(\theta_\perp(\theta_\perp'\Sigma\theta_\perp)^{-1}\theta_\perp'\Lambda_2 \int_w^1 [\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \\
& \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')'(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp')'
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \; & vec(\int_w^1 [\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')' \\
& ([\int_w^1 (\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& (\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))'dt]^{-1} \otimes I_{k-r}) \\
& vec(\int_w^1 [\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + \gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)') \\
\Rightarrow \; & vec(\int_w^1 [W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')' \\
& ([\int_w^1 (W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& (W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))'dt]^{-1} \otimes I_{k-r})
\end{aligned}
$$

$$vec(\int_w^1 [W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')$$

So, that the objective function shows a limiting behavior like,

$$\begin{aligned}
& G(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\theta}) \\
\Rightarrow \quad & vec((\int_0^w W_1 dW_1')')'((\int_0^w W_1 W_1')^{-1} \otimes I_{k-r})vec((\int_0^w W_1 dW_1')') \\
& + vec(\int_w^1 [W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)')' \\
& ([\int_w^1 (W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)) \\
& (W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t))' dt]^{-1} \otimes I_{k-r}) \\
& vec(\int_w^1 [W_1(w) + \Lambda_1^{-1}\alpha_\perp'\beta_\perp(\gamma_\perp'\beta_\perp)^{-1}\gamma_\perp'\gamma_\perp(\theta_\perp'\gamma_\perp)^{-1}\Lambda_2 W_1(t)]dW_1(t)') \\
\\
\Rightarrow \quad & vec((\int_0^w W_1 dW_1')')'((\int_0^w W_1 W_1')^{-1} \otimes I_{k-r})vec((\int_0^w W_1 dW_1')') \\
& + vec(\int_w^1 [\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + W_1(t)]dW_1(t)')' \\
& ([\int_w^1 (\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + W_1(t)) \\
& (\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + W_1(t))' dt]^{-1} \otimes I_{k-r}) \\
& vec(\int_w^1 [\Lambda_2^{-1}\theta_\perp'\gamma_\perp(\gamma_\perp'\gamma_\perp)^{-1}\gamma_\perp'\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\Lambda_1 W_1(w) + W_1(t)]dW_1(t)').
\end{aligned}$$

References

[1] Engle, R.F., and C.W.J. Granger, 1987, Co-integration and Error Correction : Representation, Estimation and Testing, *Econometrica,* 55, 251-276

[2] Hamilton, J.D., 1994, *Times Series Analysis,* Princeton University Press, Princeton, New Jersey

[3] Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica,* 50, 1029-1054

[4] Johansen, S., 1991, Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models, *Econometrica,* 59, 1551-1581

[5] Kleibergen, F., and H.K. van Dijk, 1994, Direct Cointegration Testing in Error Correction Models, *Journal of Econometrics,* 40, 63-86

[6] Kleibergen, F. and H.K. van Dijk, 1994, On the Shape of the Likelihood/Posterior in Cointegration Models, *Econometric Theory,* 10, 514-551

[7] Kleibergen, F., and P.H. Franses, *Direct Cointegration Testing in Periodic Vector Autoregressive Models,* Tinbergen Institute Research papers, 95-178

[8] Newey, W.K., and K.D. West, 1987, A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica,* 55, 703-708

[9] Phillips, P.C.B., 1983, Exact Small Sample Theory in the Simultaneous Equations Model, in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics,* Vol. 1, North-Holland Publishing Co., Amsterdam

[10] Phillips, P.C.B., 1987, Time Series Regression with a Unit Root, *Econometrica,* 55, 277-301

[11] Phillips, P.C.B., 1991, Optimal Inference in Cointegrated Systems, *Econometrica,* 59, 283-307

[12] Phillips, P.C.B., 1994, Some Exact Distribution Theory for Maximum Likelihood Estimators of Cointegrating Coefficients in Error Correction Models, *Econometrica,* 62, 73-95

[13] Phillips, P.C.B., and P. Perron, 1988, Testing for a Unit Root in Time Series Regression, *Biometrika,* 75, 335-346

[14] Phillips, P.C.B., and V. Solo, 1992, Asymptotics for Linear Processes, *Annals of Statistics,* 20, 971-1001

[15] White, H., 1980, A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica,* 48, 817-838