UPPSALA UNIVERSITET

Working Paper 2018:2

Department of Statistics

Estimating a VECM for a Small Open Economy

Sebastian Ankargren and Johan Lyhagen

Working Paper 2018:2
August 2018
Department of Statistics
Uppsala University
Box 513
SE-751 20 UPPSALA
SWEDEN

Title: Estimating a VECM for a small open economy
Author: Sebastian Ankargren & Johan Lyhagen
E-mail: sebastian.ankargren@statistik.uu.se

Estimating a VECM for a small open economy

Sebastian Ankargren* and Johan Lyhagen†

Department of Statistics, Uppsala University

Abstract

One of the most popular ways to model macro economic variables is by the vector error correction model (VECM). Besides forecasting and testing of hypotheses, the VECM is often used for calculating impulse responses, which describe how shocks today affect the variables in the future. In economic theory, a small open economy denotes the economy of a country which is too small to influence the surrounding world. The surrounding world can, for this reason, be seen as exogenous relative to the economy of this small open economy. The main contribution of this paper is the proposal of how to estimate a VECM with exogeneity restrictions on both the short-run dynamics and the short-run adjustment parameters between small open economies and the surrounding world. A Monte Carlo simulation of impulse responses shows that the proposed model is considerably more efficient compared to models fully or partially ignoring exogeneity. It is also shown that the empirical size when testing for the number of long-run relations is closer to the nominal size. Using two Swedish macroeconomic data sets the proposed method is applied to estimate the models under weak exogeneity and Granger non-causality, respectively. We find for some variables large deviances in impulse responses between our proposed model incorporating both types of restrictions and models using none or only one type of restriction, thus illustrating the need for imposing the full set of restrictions instead of settling for just one.

Keywords: VECM; Impulse responses; Small open economy; Exogeneity; Cointegration.

1 Introduction

A small open economy is an economy considered too small to influence the surrounding world, for example in terms of world prices or interest rates. This label applies to most countries today, except for the leading economies such as the United States and China, or regions like the EU in which the countries grouped together no longer constitute a small open economy. Hence, there is a great interest in developing economic theories that apply to small

*Corresponding author. E-mail address: sebastian.ankargren@statistics.uu.se

†E-mail address: johan.lyhagen@statistics.uu.se, financial support from Jan Wallanders and Tom Hedelius Stiftelser, P2016-0293:1, gratefully acknowledged.

open economies. Statistics has played a crucial role in economics since the seminal paper of Haavelmo (1944), which lay the statistical foundation for applied macro economics and theoretically justified the work of the Cowles commission. The main focus when modeling macro economic relations was, at the time, on large systems of equations and the majority of the work in macro econometrics focused on related issues, including identification, endogeneity, system estimation, etc. Focus was shifted when Sims (1980) criticized large-scale econometric models and popularized the VAR approach. An important contribution of the Sims (1980) paper is the illustration of the usefuleness of impulse responses.

Engle and Granger (1987) introduced cointegration with the possibility of modeling economic equilibriums. The main breakthrough has been the Johansen approach to cointegration (see e.g. Johansen 1988, 1991, 1995) which merged the VAR model with cointegration. This made it possible to test economic theories through the cointegrating relations and stochastic trends, but also for policy evaluations using impulse responses. With the Johansen approach it is possible to estimate and test restrictions on the cointegrating, or equilibrium, relations, as well as on the adjustment parameters, which describe how the system moves towards equilibrium. The interpretation of a restriction on the adjustment parameters is that the feedback from deviations from the long-run equilibrium—the cointegrating relation—to the corresponding dependent variable is constrained. This is the concept of weak exogeneity applied to cointegration, see Engle, Hendry, and Richard (1983).

Johansen proposed the use of reduced rank regression to estimate the parameters of the cointegrating relations and the adjustment parameters. To solve the problem with short-run dynamics, the first step is to use the Frisch-Waugh-Lovell theorem to concentrate out the short-run dynamics. When employing the model for a small open economy it is customary to have two sets of variables. The first set is the variables of interest for the economy of that country (e.g. GDP, exports, imports and inflation) and the other set contains foreign variables (such as foreign GDP and interest rates). It is important to notice that the concept of a small open economy implies no feedback from the small economy to the foreign, as these can be seen as price takers. Currently, if this is taken into account in the model it is usually accomplished by restricting the appropriate adjustment parameters to zero, but oftentimes it is simply ignored. In possibly restricting the short-run parameters as well, a problem arises as they cannot be concentrated out as in Johansen's approach. Hence, restrictions on the short-run parameters are very rare, but can be made using the results of Lütkepohl (2005) implemented in the software jMulti. The drawback is that it is not possible to simultaneously have restrictions on the short-run dynamics and the adjustment parameters. The purpose of this paper is to propose an estimation procedure which can accommodate restrictions on the short-run dynamics, the adjustment parameters and the cointegrating relations simultaneously. The procedure is based on the results of Boswijk (1995) and Groen and Kleibergen (2003).

The paper is organized as follows. The next section introduces the model and the main restrictions of interest as well as the estimation procedure. Section 3 analyzes the performance of imposing the restrictions by Monte Carlo methods while an empirical example is the topic of Section 4. A conclusion ends the paper.

2 The model and estimation

The vector error correction model, VECM, for the $k \times 1$ vector y_t can be written, ignoring deterministic terms, as

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta y_{t-i} + \varepsilon_t, \quad t = 1, \ldots, T \tag{1}$$

where α and β are full column rank matrices of size $k \times r$, Γ_i are $k \times k$ matrices containing the short-run dynamics parameters, ε_t is a $k \times 1$ vector of white noise disturbances with covariance matrix Ω and $\Delta y_t = y_t - y_{t-1}$. Following Johansen (1995), we assume that y_t is integrated of order one, meaning that it is dominated by a random walk, its first difference Δy_t is stationary and the cointegrating relation $\beta' y_t$ is stationary (and can be thought of as economic equilibriums). The matrix α describes how quickly deviations from the cointegrating relations vanish.

Next, let $y_t' = \left[y_{f,t}', y_{d,t}'\right]'$ where $y_{f,t}$ is the set of s (exogenous) variables from the foreign economy and $y_{d,t}$ the $k-s$ (endogenous) domestic variables. A small open economy paradigm implies that $\Delta y_{d,t-1}, \ldots, \Delta y_{d,t-p}$ should not affect $\Delta y_{f,t}$ and as such motivates the restricted model

$$\begin{aligned} \begin{bmatrix} \Delta y_{f,t} \\ \Delta y_{d,t} \end{bmatrix} &= \alpha \beta' \begin{bmatrix} y_{f,t-1} \\ y_{d,t-1} \end{bmatrix} + \sum_{i=1}^{p} \Gamma_i \begin{bmatrix} \Delta y_{f,t-i} \\ \Delta y_{d,t-i} \end{bmatrix} + \varepsilon_t \\ &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \beta_{11}' & \beta_{21}' \\ \beta_{12}' & \beta_{22}' \end{bmatrix} \begin{bmatrix} y_{f,t-1} \\ y_{d,t-1} \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \Gamma_{11i} & 0 \\ \Gamma_{21i} & \Gamma_{22i} \end{bmatrix} \begin{bmatrix} \Delta y_{f,t-i} \\ \Delta y_{d,t-i} \end{bmatrix} + \varepsilon_t. \end{aligned} \tag{2}$$

To estimate the model, we first note that as $\Delta y_{f,t-i}$ influences all left-hand side variables and are without restrictions we can use the Frisch-Waugh-Lovell theorem to concentrate them out. Let $\Delta \tilde{y}_t$, $\tilde{y}_{t-1}$, $\Delta \tilde{y}_{d,t-i}$ and $\tilde{\varepsilon}_t$ denote the remaining variables with $\Delta y_{f,t-i}$ partialed out. Stacking the observations yields the model in matrix form

$$\tilde{Y} = \tilde{Y}_{-1} \Pi + \tilde{\varepsilon} \tag{3}$$

where

$$\Pi = \begin{bmatrix} \multicolumn{2}{c}{\beta \alpha'} \\ 0' & \Gamma_{221}' \\ \vdots & \vdots \\ 0' & \Gamma_{22p}' \end{bmatrix} \tag{4}$$

and

$$\tilde{Y} = \begin{bmatrix} \Delta\tilde{y}_1' \\ \Delta\tilde{y}_2' \\ \vdots \\ \Delta\tilde{y}_T' \end{bmatrix}, \tilde{Y}_{-1} = \begin{bmatrix} \tilde{y}_0' & \Delta\tilde{y}_{d,-1}' & \cdots & \Delta\tilde{y}_{d,-p}' \\ \tilde{y}_1' & \Delta\tilde{y}_{d,0}' & & \Delta\tilde{y}_{d,-p+1}' \\ \vdots & & \ddots & \\ \tilde{y}_{T-1}' & \Delta\tilde{y}_{d,T-1}' & & \Delta\tilde{y}_{d,T-p}' \end{bmatrix}. \tag{5}$$

The concentrated log-likelihood is, bar a constant,

$$\ell(\Pi) = -\frac{T}{2}|\Omega| - \frac{1}{2}\operatorname{vec}(\tilde{Y} - \tilde{Y}_{-1}\Pi)'(\Omega^{-1} \otimes I_T)\operatorname{vec}(\tilde{Y} - \tilde{Y}_{-1}\Pi). \tag{6}$$

The second term in the above display can be rewritten into

$$G(\Pi, \Omega) = \operatorname{vec}\left[\tilde{Y}_{-1}'\left(\tilde{Y} - \tilde{Y}_{-1}\Pi\right)\right]'\left[\Omega \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]^{-1}\operatorname{vec}\left[\tilde{Y}_{-1}'\left(\tilde{Y} - \tilde{Y}_{-1}\Pi\right)\right], \tag{7}$$

where maximization of (6) is equivalent to minimization of (7). For the time being, we assume Ω is known. To accommodate minimization of (7), we first rewrite $\operatorname{vec}\left[\tilde{Y}_{-1}'\left(\tilde{Y} - \tilde{Y}_{-1}\Pi\right)\right]$ as

$$\operatorname{vec}\left[\tilde{Y}_{-1}'\left(\tilde{Y} - \tilde{Y}_{-1}\Pi\right)\right] = \operatorname{vec}\left(\tilde{Y}_{-1}'\tilde{Y}\right) - \operatorname{vec}\left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\Pi\right) \tag{8}$$

$$= \operatorname{vec}\left(\tilde{Y}_{-1}'\tilde{Y}\right) - F\pi, \tag{9}$$

where $F = \left(I_k \otimes \tilde{Y}_{-1}'\tilde{Y}_{-1}\right)$ and $\pi = \operatorname{vec}(\Pi)$. The underlying idea for the estimation procedure is that $F\pi$ can, to begin with, be written in two different ways, each conditional on one of α and β. To fix ideas, first let

$$\pi_\beta = \begin{bmatrix} \operatorname{vec}(\beta') \\ \operatorname{vec}(\Gamma_{221}) \\ \vdots \\ \operatorname{vec}(\Gamma_{22p}) \end{bmatrix}, \quad \pi_\alpha = \begin{bmatrix} \operatorname{vec}(\alpha) \\ \operatorname{vec}(\Gamma_{221}) \\ \vdots \\ \operatorname{vec}(\Gamma_{22p}) \end{bmatrix}. \tag{10}$$

Then, the generic $F\pi$ can be written as

$$F\pi = F_\beta(\alpha)\pi_\beta \tag{11}$$

$$= F_\alpha(\beta)\pi_\alpha \tag{12}$$

where

$$F_\beta(\alpha) = \left(I_k \otimes \tilde{Y}_{-1}'\tilde{Y}_{-1}\right) K_{k,k+p(k-s)} \begin{bmatrix} I_k \otimes \underset{k\times r}{\alpha} & \underset{k^2\times p(k-s)^2}{0} \\ \underset{pk(k-s)\times kr}{0} & I_{p(k-s)} \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s} \end{bmatrix} \end{bmatrix} \tag{13}$$

and

$$F_\alpha(\beta) = \left(I \otimes \tilde{Y}_{-1}'\tilde{Y}_{-1}\right) K_{k,k+p(k-s)} \begin{bmatrix} \underset{k\times r}{\beta} \otimes I_k & \underset{k^2\times p(k-s)^2}{0} \\ \underset{pk(k-s)\times kr}{0} & I_{p(k-s)} \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s} \end{bmatrix} \end{bmatrix}. \tag{14}$$

We are here using $K_{m,n}$ to denote the commutation matrix for an $m \times n$ matrix A defined by $K_{m,n} \operatorname{vec}(A) = A'$ (Magnus and Neudecker, 1979).

As is clear from (10), estimating π_α and π_β means estimating the full $k\times r$ matrices α and β without restrictions. However, when the model is used for modeling a small open economy where exogenous, large-economy variables are also included such an approach may not be reasonable. If α and β are left unrestricted then the model will allow for the r cointegrating relations $\beta' y_{t-1}$ to enter all equations in the system, including the equations for the foreign variables. As a small open economy should not be able to have this influence, shutting down such a connection is in many modeling situations warranted. We consider two approaches to enforce the small open economy property into our model.

2.1 Weak exogeneity

The first approach assumes the foreign variables to be weakly exogenous for the cointegrating vectors β' such that $\alpha = (0_{s\times r}', \alpha_2')'$, where α_2 is $(k-s)\times r$. Weak exogeneity naturally leads to the notion of conditional and marginal models (Johansen, 1995) from which the full model can be reconstructed (see also the discussion in Jacobs and Wallis, 2010). Such a restriction, together with the restriction on the short-run parameters Γ made in (2), effectively eliminates feedback from the domestic variables to the foreign. We let the parameter vector π_α under this paradigm be denoted by $\pi_\alpha^{(WE)}$ where

$$\pi_\alpha^{(WE)} = \begin{bmatrix} \operatorname{vec}(\alpha_2) \\ \operatorname{vec}(\Gamma_{221}) \\ \vdots \\ \operatorname{vec}(\Gamma_{22p}) \end{bmatrix} \tag{15}$$

which differs from π_α in that $\pi_\alpha^{(WE)}$ only includes α_2 and not $\alpha_1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} \end{bmatrix}$ (which is now fixed to be 0). The corresponding F matrix in (9) is given by

$$F_\alpha^{(WE)}(\beta) = F_\alpha(\beta) \begin{bmatrix} I_r \otimes \begin{bmatrix} 0_{s\times(k-s)} \\ I_{k-s} \end{bmatrix} & 0_{rk\times p(k-s)^2} \\ 0_{p(k-s)^2\times r(k-s)} & I_{p(k-s)^2} \end{bmatrix}. \tag{16}$$

2.2 Granger non-causality

The second approach (which we refer to as Granger non-causality, see for example Toda and Phillips, 1993) to eliminating feedback from domestic to foreign variables is to allow for some cointegrating relations to enter the foreign equations, but only those which solely involve foreign variables. In our application later, we will revisit a model for the Swedish economy in which there are four domestic and three foreign variables. In this case, two cointegrating relations among all variables exclusively enter the domestic part of the system and one cointegrating relation involving only foreign variables is allowed to enter all equations.

Mathematically, this imposes restrictions not only on α but also on β. To see this, suppose that there are r_1 foreign and r_2 domestic cointegrating relations with $r_1 + r_2 = r$. If we partition α and β into blocks with r_1 and r_2 columns and s and $k-s$ rows we have

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}. \tag{17}$$

This second approach then means to let $\alpha\beta'$ be

$$\alpha\beta' = \begin{pmatrix} \alpha_{11} & 0_{s\times r_2} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \beta'_{11} & 0_{r_1\times k-s} \\ \beta'_{12} & \beta'_{22} \end{pmatrix}. \tag{18}$$

Such restrictions, coupled with the zero constraints on the Γ_i matrices, enforce Granger non-causality of domestic variables on the foreign as also discussed by Toda and Phillips (1993). In fact, Granger non-causality can be achieved by the less restrictive assumption $\alpha_{11}\beta'_{11} + \alpha_{12}\beta'_{21} = 0$. However, this enforces a non-linear restriction which complicates the statistical analysis dramatically as Toda and Phillips (1993) demonstrate. We therefore maintain the above assumption for the block-restricted model.

Estimation is carried out as in the preceding case but now using F matrices defined by

$$\begin{aligned} F_\beta^{(GN)}(\alpha) &= \left(I_k \otimes \tilde{Y}'_{-1}\tilde{Y}_{-1}\right) K_{k,k+p(k-s)} \\ &\times \left[\begin{array}{c|c|c} \begin{bmatrix} I_s \\ 0_{k-s\times s}\end{bmatrix} \otimes \alpha & \begin{matrix} 0_{ks\times r_2(k-s)} \\ I_{k-s}\otimes \alpha \begin{bmatrix} 0_{r_1\times r_2} \\ I_{r_2}\end{bmatrix}\end{matrix} & 0_{r(k-s)\times p(k-s)^2} \\ \hline 0_{pk(k-s)\times rs} & 0_{pk(k-s)\times r_2(k-s)} & I_{p(k-s)} \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s}\end{bmatrix} \end{array}\right] \\ F_\alpha^{(GN)}(\beta) &= F_\alpha(\beta) \left[\begin{array}{c|c|c} K_{r,k}\begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} \\ 0_{r_2\times r_1}\end{bmatrix} \\ 0_{r(k-s)\times r_1 s}\end{bmatrix} K_{s,r_1} & I_r \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s}\end{bmatrix} & 0_{rk\times p(k-s)^2} \\ \hline 0_{p(k-s)^2\times r_1 s} & 0_{p(k-s)^2\times r(k-s)} & I_{p(k-s)^2} \end{array}\right] \end{aligned} \tag{19}$$

where the parameter vectors are

$$\pi_\beta^{(GN)} = \begin{bmatrix} \mathrm{vec}(\beta'_1) \\ \mathrm{vec}(\beta'_{22}) \\ \mathrm{vec}(\Gamma) \end{bmatrix}, \quad \pi_\alpha^{(GN)} = \begin{bmatrix} \mathrm{vec}(\alpha_{11}) \\ \mathrm{vec}(\alpha_2) \\ \mathrm{vec}(\Gamma) \end{bmatrix}. \tag{20}$$

2.3 Estimation

By substituting $\text{vec}\left[\tilde{Y}_{-1}'\left(\tilde{Y} - \tilde{Y}_{-1}\Pi\right)\right]$ for (9) and $F\pi$ for the desired choice of F and π it is possible to solve for π_β and π_α, respectively. The solutions are

$$\hat{\pi}_\beta(\alpha, \Omega) = \left\{F_\beta(\alpha)'\left[\Omega \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]^{-1} F_\beta(\alpha)\right\}^{-1} F_\beta(\alpha)'\left[\Omega \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]^{-1} \text{vec}\left(\tilde{Y}_{-1}'\tilde{Y}\right) \tag{21}$$

$$\hat{\pi}_\alpha(\beta, \Omega) = \left\{F_\alpha(\beta)'\left[\Omega \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]^{-1} F_\alpha(\beta)\right\}^{-1} F_\alpha(\beta)'\left[\Omega \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]^{-1} \text{vec}\left(\tilde{Y}_{-1}'\tilde{Y}\right). \tag{22}$$

Enforcing weak exogeneity or Granger non-causality simply amounts to substituting the F matrices yielding the unrestricted estimates with $F^{(WE)}$ or $F^{(GN)}$ in (21)–(22).

Finally, rarely ever is Ω a known matrix and it too must be estimated. To facilitate this, we use the conditional maximum likelihood estimator of Ω conditional on Π given by

$$\hat{\Omega}(\Pi) = \frac{1}{T}(Y - Y_{-1}\Pi)'(Y - Y_{-1}\Pi). \tag{23}$$

Evidently, there is a circular dependence in the equations which implicitly suggests an iterative estimation procedure. This iterative procedure is as follows:

1. Estimate Ω, α in an unrestricted VECM
2. Estimate π_β using $\hat{\pi}_\beta(\hat{\alpha}, \hat{\Omega})$ in (21)
3. Estimate π_α using $\hat{\pi}_\alpha(\hat{\beta}, \hat{\Omega})$ in (22)
4. Estimate Ω using $\hat{\Omega}(\hat{\Pi})$ in (23)
5. Iterate 2–4 until convergence

Such a switching algorithm has previously been applied in the cointegration literature by e.g. Johansen and Juselius (1992, 1994); Groen and Kleibergen (2003); Boswijk and Doornik (2004). While none of the previous studies have proven that the algorithm converges to a global maximum, each step is non-decreasing in the likelihood and generally works very well. The asymptotic properties of the estimator based on (19)–(20) are established in the following proposition in which we restrict ourselves to the case without deterministic terms for simplicity, but without loss of generality.

Proposition 1. *Assume that the model is*

$$\begin{bmatrix} \Delta y_{f,t} \\ \Delta y_{d,t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & 0 \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \beta_{11}' & 0 \\ \beta_{12}' & \beta_{22}' \end{bmatrix} \begin{bmatrix} y_{f,t-1} \\ y_{d,t-1} \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \Gamma_{11i} & 0 \\ \Gamma_{21i} & \Gamma_{22i} \end{bmatrix} \begin{bmatrix} \Delta y_{f,t-i} \\ \Delta y_{d,t-i} \end{bmatrix} + \varepsilon_t, \tag{24}$$

where

1. *the error sequence* $\{\varepsilon_t\}$ *is such that: 1) it is a martingale difference sequence satisfying* $E(\varepsilon_t) = 0$ *and* $E(\varepsilon_t \varepsilon_{t-\tau}) = 0$ *for* $\tau > 0$, *2) the strong law of large numbers applies so that* $T^{-1} \sum_{t=1}^{T} E(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}) \xrightarrow{a.s.} \Omega$ *where* $\mathcal{F}_{t-1}$ *is the filtration up to time* $t-1$,

2. α *and* β $(k \times r)$ *are full rank,*

3. $y_t \sim I(1)$ *so that* $\alpha_\perp' \Gamma \beta_\perp$ *is non-singular.*

Using (19)–(20) *for estimation, the asymptotic distribution of the long-run parameters is*

$$T\left(\begin{bmatrix}\mathrm{vec}(\hat{\beta}_1')\\ \mathrm{vec}(\hat{\beta}_{22}')\end{bmatrix} - \begin{bmatrix}\mathrm{vec}(\beta_1')\\ \mathrm{vec}(\beta_{22}')\end{bmatrix}\right) \xrightarrow{d} \tag{25}$$

$$\begin{bmatrix} \int G_{k,1}(u)G_{k,1}(u)'du \otimes \alpha'\Omega^{-1}\alpha & \int G_{k,1}(u)G_{k,2}(u)'du \otimes \alpha'\Omega^{-1}\alpha_{\cdot 2} \\ \int G_{k,2}(u)G_{k,1}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha & \int G_{k,2}(u)G_{k,2}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha_{\cdot 2} \end{bmatrix}^{-1} \tag{26}$$

$$\times \begin{bmatrix} \mathrm{vec}\left(\alpha'\Omega^{-1} \int dW_k G_{k,1}'\right) \\ \mathrm{vec}\left(\alpha_{\cdot 2}'\Omega^{-1} \int dW_k G_{k,2}'\right) \end{bmatrix}, \tag{27}$$

where $G_{k,1}(u)$ *and* $G_{k,2}(u)$ *denote the first* s *and last* $k-s$ *elements of* $G_k(u) = CW_k(u)$, *respectively; here,* $C = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'$, $\Gamma = I_k - \sum_{i=1}^{p-1}\Gamma_i$ *and* $W_k(u)$ *is a* k*-dimensional Brownian motion with covariance matrix* Ω, *where also* $\alpha_{\cdot 2}' = \begin{bmatrix}\alpha_{12}' & \alpha_{22}'\end{bmatrix}$.

Moreover, the asymptotic distribution of the short-run parameters estimated based on (19)–(20) *is*

$$\sqrt{T}\left(\hat{\pi}_\alpha^{(GN)} - \pi_\alpha^{(GN)}\right) \xrightarrow{d} \tag{28}$$

$$N_{r_1 s + r(k-s) + p(k-s)^2}\left(0, \begin{bmatrix} \Sigma_{\beta\beta,11} \otimes (\Omega^{-1})_{11} & \Sigma_{\beta\beta,1\cdot} \otimes (\Omega^{-1})_{12} & \Sigma_{\beta 0,1\cdot} \otimes (\Omega^{-1})_{12} \\ \Sigma_{\beta\beta,\cdot 1} \otimes (\Omega^{-1})_{21} & \Sigma_{\beta\beta} \otimes (\Omega^{-1})_{22} & \Sigma_{\beta 0} \otimes (\Omega^{-1})_{22} \\ \Sigma_{0\beta,\cdot 1} \otimes (\Omega^{-1})_{21} & \Sigma_{0\beta} \otimes (\Omega^{-1})_{22} & \Sigma_{00} \otimes (\Omega^{-1})_{22} \end{bmatrix}^{-1}\right), \tag{29}$$

where $\Sigma_{\beta\beta} = \mathrm{plim}\, T^{-1}\sum_{t=1}^{T}\beta' y_t y_t'\beta$, $\Sigma_{\beta 0} = \mathrm{plim}\, T^{-1}\sum_{t=1}^{T}\beta' y_{t-1}\Delta y_t'$ *and* $\Sigma_{00} = \mathrm{plim}\, T^{-1}\sum_{t=1}\Delta y_t \Delta y_t'$; $\Sigma_{\beta\beta,11}$ *is the* $r_1 \times r_1$ *upper left block of* $\Sigma_{\beta\beta}$, *whereas* $\Sigma_{\beta\beta,\cdot 1}$ *and* $\Sigma_{\beta 0,\cdot 1}$ *refer to the first* r_1 *columns of the corresponding matrix. Similarly,* $(\Omega^{-1})_{ij}$ *is defined to be the* (i,j)*th block of* Ω^{-1} *(whose blocks have* s *or* $k-s$ *rows and/or columns).*

The proof is placed in the Appendix.

Remark 1. If the model is estimated under the weak exogeneity restriction $r_1 = 0$ and so

$$T\,\mathrm{vec}\left(\hat{\beta}^{(WE)'} - \beta^{(WE)'}\right) \tag{30}$$

$$\xrightarrow{d} \left(\int G_k(u)G_k(u)'du \otimes \alpha'\Omega^{-1}\alpha\right)^{-1} \mathrm{vec}\left(\alpha'\Omega^{-1}\int dW_k G_k'\right) \tag{31}$$

$$= \mathrm{vec}\left[\int dV_k G_k' \left(\int G_k(u)G_k(u)'du\right)^{-1}\right] \tag{32}$$

where $V_k(u) = (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}W_k(u)$. The latter form exactly mirrors the result by Johansen (1995).

Furthermore,

$$\sqrt{T}\left(\hat{\pi}_\alpha^{(WE)} - \pi_\alpha^{(WE)}\right) \xrightarrow{d} N_{r(k-s)+p(k-s)^2}\left(0, \Sigma^{-1} \otimes (\Omega^{-1})_{22}^{-1}\right), \tag{33}$$

where

$$\Sigma = \begin{bmatrix} \Sigma_{\beta\beta} & \Sigma_{\beta 0} \\ \Sigma_{0\beta} & \Sigma_{00} \end{bmatrix}. \tag{34}$$

Remark 2. If there are no foreign variables such that $r_1 = s = 0$, then

$$\sqrt{T}\left(\hat{\pi}_\alpha - \pi_\alpha\right) \xrightarrow{d} N_{rk+pk^2}\left(0, \Sigma^{-1} \otimes \Omega\right), \tag{35}$$

and we obtain the same asymptotic distribution as derived by Lütkepohl (2005).

3 Monte Carlo simulation

To analyze the small-sample properties we generate data according to (2). We generate five variables, whereof the first two exogenous, with two cointegrating relations and three lags, i.e. $n = 5$, $s = 2$, $p = 3$ for the sample sizes $T = 100, 125, 150, \ldots, 500$. The parameter values are randomly chosen prior to the simulation and not altered. Figure 6 in Appendix A displays the eigenvalues for the model, which summarize its dynamic properties. The number of replicates is 5,000. As one main purpose of many macro economic modeling exercises is to estimate impulse responses we mainly evaluate using deviations from the true impulse responses, in form of mean squared error (MSE). The size property of testing the null of two cointegrating relations and the power of testing the null of one relationship is also investigated. The models we compare are the following VECM i) unrestricted, ii) restrictions on α (weak exogeneity), iii) restrictions on Γ_i (short-run restrictions), and iv) restrictions on both α and Γ_i.

In Figure 1 the results for the trace test for cointegrating rank is shown. The two models we compare are the standard VECM and a VECM with restrictions on the short-run dynamics. The leftmost figure shows the rejection rates for the true null of a rank equal to two. The middle and rightmost figures display the size-adjusted and raw powers, respectively, for the case of a rank of two. The size adjustment is carried out by simulating a model with rank one and adjusting size accordingly. The result is that the VECM model with short-run restrictions has slightly better nominal size. The higher nominal size for the unrestricted VECM carries over to the higher raw power compared to the restriced VECM. After size adjustment, the power is very similar but with a very minor power advantage for the restricted VECM.

In Figures 2 and 3 we display MSE for impulse responses as a function of impulse response horizon and sample size respectively. The sample size in Figure 2 is 100 and it can be seen that the weakly exogenous VECM with short-run restrictions always has the lowest MSE

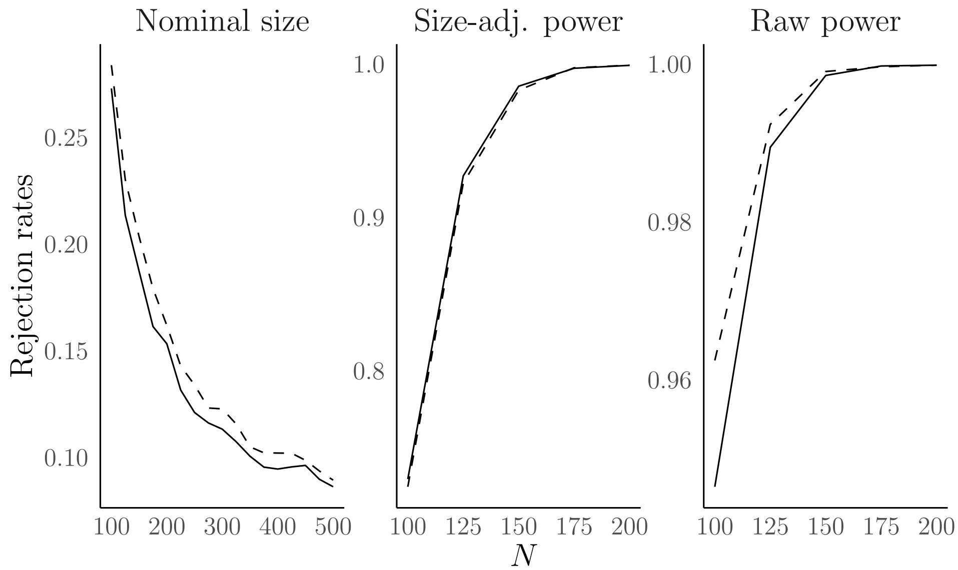

Figure 1: Nominal size, size-adjusted power and raw power using standard VECM (——) and VECM with short-run restrictions and weak exogeneity (········).

while the standard VECM has the largest. Sometimes the weakly exogenous VECM is better than the VECM with short-run restrictions, as in the top two figures, and sometimes it is the other way round as in the bottom two figures. Interestingly, the MSE always increases for the standard VECM when increasing the impulse response horizon, while this is not always true for the other models. The interpretation of this is that the restrictions are important to model the long-run relations, although they are not formally defined as long-run parameters.

Figure 3 show MSE as a function of sample size. As in the previous figure, the standard VECM has the largest MSE while the weakly exogenous VECM with short-run restrictions has the smallest. Also, as above, the ordering in terms of MSE between the VECM with only short-run dynamics versus only weakly exogenous restrictions is inconclusive, but always between the standard VECM and the model with restrictions on both the short-run parameters and weak exogeneity restrictions. As can be expected, as all models nest the true model, the MSE decreases with increased sample size. Sometimes the relative difference is small, as in the top-right figure, and in other cases the difference is large, as in the bottom figures. Quite often, the differences are substantial.

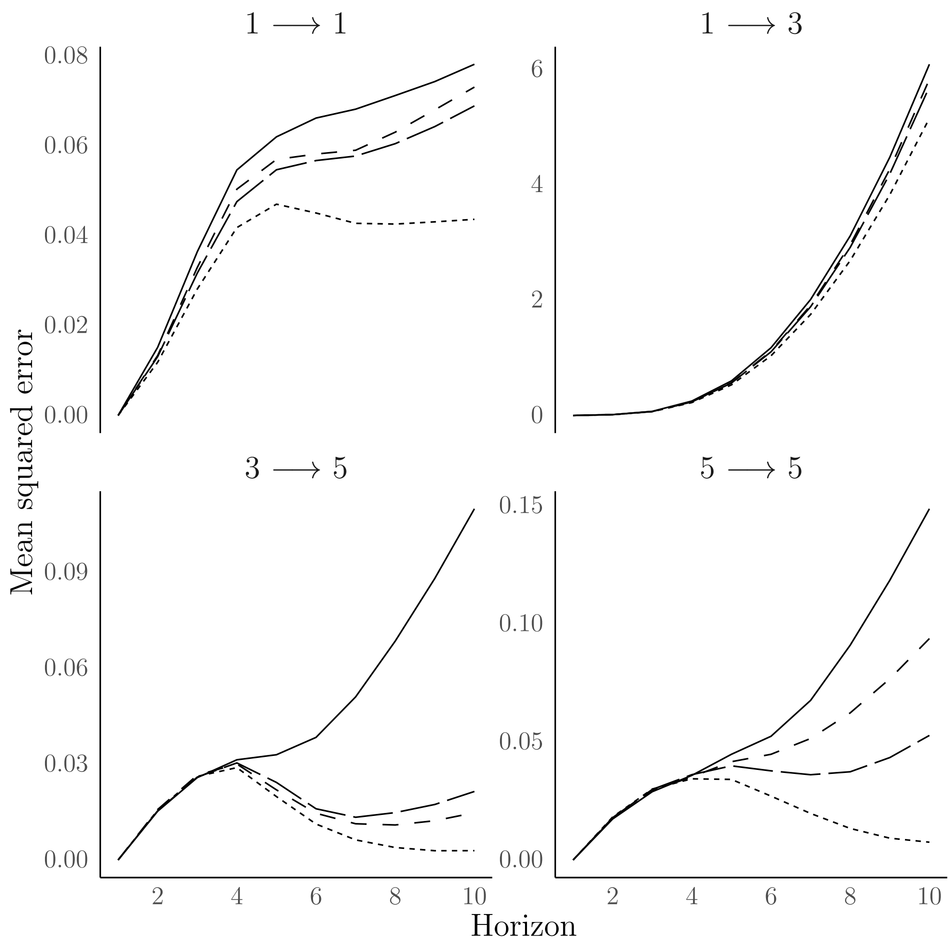

Figure 2: Mean squared error of impulse responses as a function of horizon, sample size $T = 100$. The four lines in the figures represent standard VECM (——), VECM with weak exogeneity (– –), VECM with short-run restrictions (- - -) and VECM with short-run restrictions and weak exogeneity (·········).

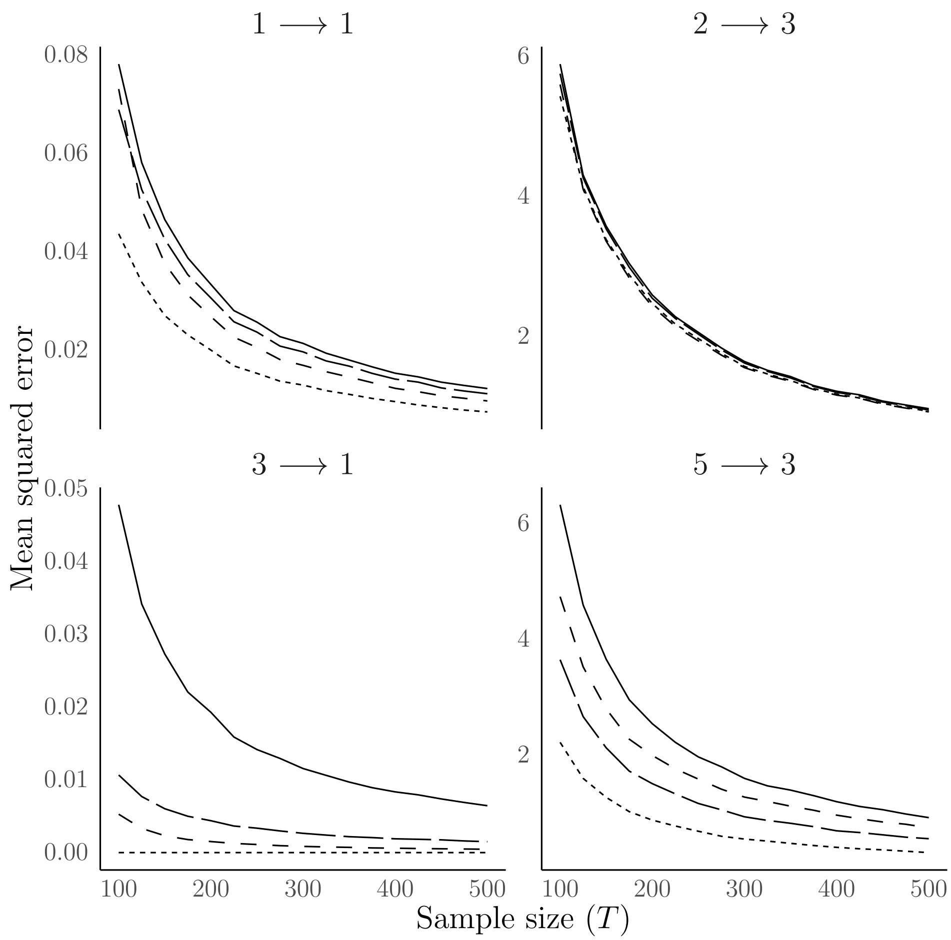

Figure 3: Mean squared error of impulse responses as a function of sample size at impulse response horizon $h = 10$. The four lines in the figures represent standard VECM (——), VECM with weak exogeneity (– –), VECM with short-run restrictions (- - -) and VECM with short-run restrictions and weak exogeneity (·········).

Table 1: Variable description for the Swedish data

Variable label	**Description**
SWEGDP*	Seasonally adjusted real GDP, in logarithms
KIX	Competitor-weighted effective exchange rate index, in logarithms
CPIX	Underlying inflation index, in logarithms
TWGDP**	Foreign GDP as weighted between the US GDP and the euro zone's GDP, in logarithms
TB	Closing yield for a 3-months treasury bill
UNEMP	Relative unemployment
Dummy	Dummy variable of one from 1991:Q4 to and including 1992:Q3, and zero otherwise

Note: KIX, TB and UNEMP are aggregated to quarterly frequencies by taking averages of the corresponding months.
*SWEGDP is sometimes referred to as GDP only for readability.
**The weights in TWGDP is 0.25 the US and 0.75 the Euro zone.

4 Empirical analysis

4.1 A weakly exogenous model for Sweden

The Swedish Ministry of Finance produces one of the most important GDP forecasts for the Swedish economy. As simple baseline models, the Ministry of Finance uses various types of vector autoregressive models, see e.g. Bjellerup and Shahnazarian (2012). We follow the same track and use a cointegrated VAR model, where the main variable of interest is the logarithm of GDP. Other variables in the model are a competitor-weighted exchange rate index, consumer price index, a foreign trade-weighted GDP (the US and the EU), interest on Swedish 3-month Treasury bills and unemployment. Similarly to Bjellerup and Shahnazarian (2012) we also have a dummy for the period 1991:Q3 to 1992:Q3. This data set has previously been used in Lyhagen, Ekberg, and Eidestedt (2015) which investigated the effect of intercept correction on forecasts of GDP. Table 2 summarizes the data, which ranges from 1988:Q1 up to 2015:Q4.[1]

The specification of the VECM closely follows Lyhagen, Ekberg, and Eidestedt (2015) with two cointegrating relations, found by using the p-values of MacKinnon, Haug, and Michelis (1999), and four lags in levels. As a measure of evaluation we use the same as in the Monte Carlo simulation above, impulse responses. The size of the model is commonly found in the literature of empirical VECM models. The number of observations, $T = 116$, is typical for this type of application.

In Figure 4 examples of the impulse responses from the model of the Swedish economy is shown (the full set of impulse responses can be found in Figure 7 in Appendix B). Similarly to the previous section we consider one standard deviation shocks without any further identi-

[1] The data constitutes an extension of the data set compared to Lyhagen, Ekberg, and Eidestedt (2015), who used data for the period 1989:Q4–2012:Q2.

fication scheme; the main purpose is to illustrate the differences obtained with changing sets of restrictions, and that message will remain under sophisticated identification approaches. Overall, imposing both short-run dynamics and weak exogeneity restrictions sometimes lead to a different picture being painted than if none or only one type of restriction is enforced.

For example, in the middle-right figure the impulse responses of a one standard deviation shock of Swedish GDP on the trade-weighted GDP is displayed (SWEGDP $\rightarrow$ TWGDP). As TWGDP is assumed to be exogenous, the VECM with both short-run and weakly exogenous restrictions yields a straight line at zero (dotted line). Only restricting the short-run dynamics (dashed line) results in a negative impact of a shock to Swedish GDP on the trade-weighted GDP. When ignoring short-run restrictions a positive effect emerges with a larger impact of the standard VECM (solid line) compared to the VECM with weakly exogenous restrictions (longdashed line). Thus, failing to cancel the channel from SWEGDP and TWGDP completely leads to the unrealistic result that Swedish GDP shocks affect trade-weighted GDP.

Switching to a one standard deviation shock in the trade-weighted GDP and its effect on Swedish GDP in the bottom-right figure (TWGDP $\rightarrow$ SWEGDP) we find an initial positive impact for all models which levels out at a positive long-term effect of around 1–1.5 and all models agree relatively well. In contrast, the effect of a one standard deviation shock of the trade-weighted GDP on underlying inflation in the bottom-left figure (TWGDP $\rightarrow$ CPIX) yields a negative impulse response for the standard VECM and the VECM with short-run dynamics restricted, whereas there are positive effects indicated by the weak exogeneity and fully-restricted models.

Shocks in exchange rates yield no effects on inflation (KIX $\rightarrow$ CPIX) according to the weak exogeneity model, but the remaining models indicate negative effects. Lastly, exchange rate shocks on itself as displayed in the top-right figure (KIX $\rightarrow$ KIX) have no long-run effects according to the models with weak exogeneity restrictions, but a persistent positive effect is found by the models without these restrictions.

According to our empirical results there are sometimes large differences depending on the restrictions imposed. The results also clearly illustrate that occasionally models with weakly exogenous restrictions seem to behave similarly with or without restricted short-run dynamics, but that is not always the case. Hence, there is a need to consider models enforcing restrictions on both the short-run adjustment parameter α, which impose weak exogeneity, as well as on the short-run dynamics. While a model with no restrictions can still be consistently estimated under the setting in Section 2, by the nature of a small open economy the act of imposing such restrictions is uncontroversial for many applications.

4.2 Modeling Sweden using Granger non-causality

Our second example originates from the work of Jacobson, Jansson, Vredin, and Warne (2001) who analyzed monetary policy and inflation forecasting for Sweden. One of the main points is that restrictions motivated by economic theory imposed on the long-run relations (i.e. the cointegrating vectors) are useful both for policy analysis as well as for forecasting. The variables used are displayed in Table 2. Jacobson, Jansson, Vredin, and Warne (2001)

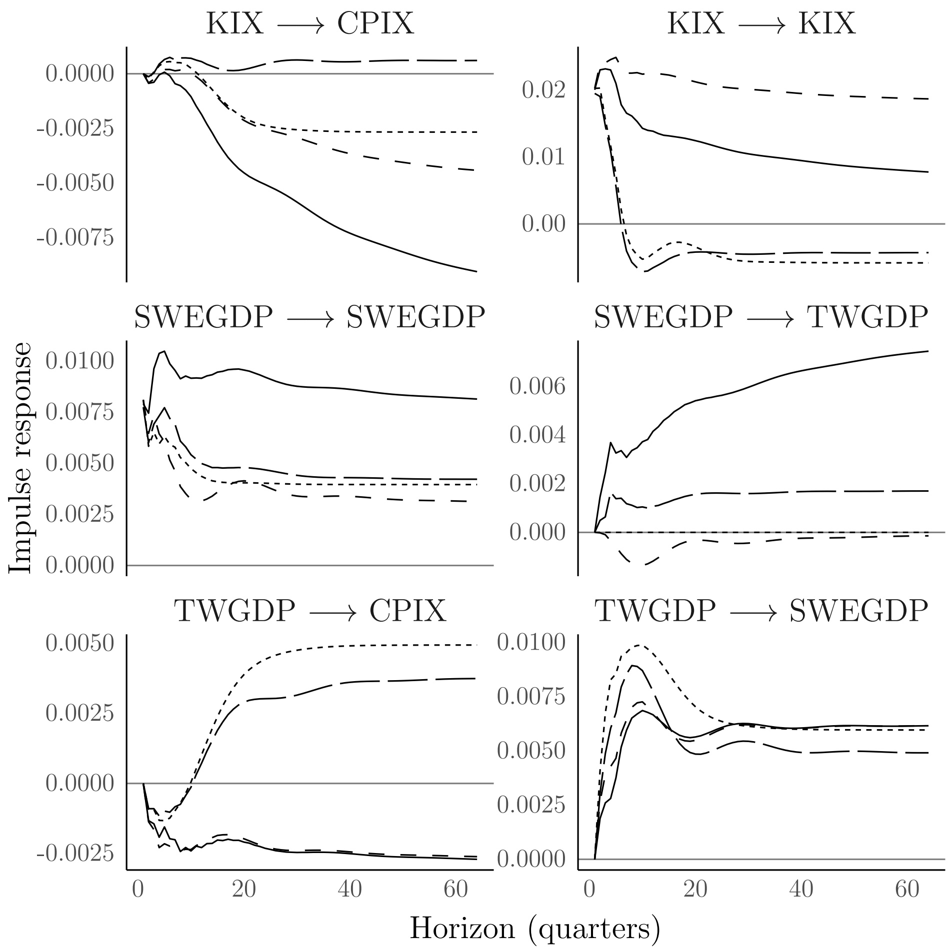

Figure 4: Examples of impulse responses to 1 SD shocks for the model of the Swedish economy (Section 4.1). The four lines in the figures represent standard VECM (——), VECM with weak exogeneity (– –), VECM with short-run restrictions (- - -) and VECM with short-run restrictions and weak exogeneity (········).

argue that there should be four stochastic trends implying three cointegrating relations as there are seven variables in total. Three of the variables are foreign and four domestic. The time period is 1972:Q2–1996:Q4 yielding a total of 99 observations. Additionally, five dummy variables for 'crashes' and 'changes in growth' capturing regime shifts in economic policy are used. The interpretation of the three cointegrating relations are that the first is a goods market equilibrium, the second is related to a financial markets equilibrium condition while the third is common trends and equilibrium conditions between the foreign variables. We use the same number of lags (four) and cointegrating vectors.

In this subsection we compare impulse responses from a standard VECM, a VECM with Granger non-causality and a VECM with Granger non-causality and long-run restrictions.[2] All impulse responses can be found in Figure 8, while a selected subset are shown in Figure 5.

As in the previous example the results differ substantially depending on if restrictions are imposed and, if so, which restrictions are imposed. For example, Swedish real output on the foreign variables as well as on all the domestic, except Swedish price levels, are very similar no matter which model is used. According to the standard VECM there is no effect from Foreign real output on Swedish price levels while for the Granger non-causal models there are positive effects. The pattern is the same for Foreign real output on Swedish interest rates but the level of the effects of the Granger non-causal models are on a lower level. The opposite is true for Foreign price levels on Swedish interest rates where the Granger non-causal models have no effect while the standard VECM has a positive effect. An interesting result is the effect of Foreign price levels on Swedish price levels where the standard model shows a large positive effect while the Granger non-causal with long-run restrictions has minor positive effects and the Granger non-causal has negative effects. The results of the restricted VECM are more reasonable as the unrestricted imply a permanent effect of the same magnitude as the initial effect. The restricted VECMs imply that the pass through is not full and that there is a change in consumption due to changed relative prices. According to a VECM with Granger non-causality and long-run restrictions the effect of Swedish price levels on Nominal exchange rates are about twice as large compared to the standard VECM, and the VECM with Granger non-causality but no long-run restrictions is just below the one with long-run restrictions. The results are the same, but on a lower level, when considering the effect of Swedish interest rates on Nominal exchange rates. The pattern is the opposite for the case of Nominal exchange rates on Swedish price levels where there are positive effects for the standard VECM but no effects for the Granger non-causal models. A shock in Swedish price levels on Swedish real output is shown to be positive for the standard VECM while negative for both restricted versions.

[2] The long-run restrictions follow Jacobson, Jansson, Vredin, and Warne (2001). Using the same ordering of variables as in Table 2, β is restricted to

$$\beta' = \begin{bmatrix} \beta_{11} & 1 & \beta_{12} & \beta_{13} & -1 & \beta_{14} & 1 \\ 0 & \beta_{21} & -1 & \beta_{22} & \beta_{23} & 1 & \beta_{24} \\ \beta_{31} & \beta_{32} & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{36}$$

Table 2: Jacobson, Jansson, Vredin, and Warne (2001) data

Variable	Description
y^f	Foreign real output, $y_t^f = 100\ln(Y_t^f)$ where Y_t^f is German real GDP in 1991 prices
p^f	Foreign price levels, $p_t^f = 100\ln(P_t^f)$, where P_t^f is the geometric sum of Sweden's 20 most important trading partners weighted by IMF's TCW index
i^f	Foreign nominal interest rate, $i_t^f = 100\ln(1 + I_t^f/100)$ where I_t^f is the German three-month treasury bills rate
y	Swedish real output, $y_t = 100\ln(Y_t)$ where Y_t is Swedish real GDP in 1991 prices
p	Swedish price levels, $p_t = 100\ln(P_t)$, where P_t is the quarterly average of Swedish CPI
i	Swedish nominal interest rate, $i_t = 100\ln(1 + I_t/100)$ where I_t is the Swedish three-month treasury bills rate
e	Nominal exchange rate, $e_t = 100\ln(S_t)$ where S_t is the geometric sum of the nominal Krona exchange rate of Sweden's top 20 trading partners using the TCW index

Note: The model also includes five dummy variables, see Jacobson, Jansson, Vredin, and Warne (2001) for details.

Overall, these results show the importance to impose theoretically motivated restrictions on the model, and not only on the long-run relations but also on the short-run dynamics.

5 Conclusions

In this paper we have proposed the use of an estimation procedure in the case of exogeneity restrictions in a vector error correction model (VECM), which naturally arise when modeling small open economies. A Monte Carlo simulation is used to show the advantages of imposing such restrictions. It is found that it is beneficial, in terms of mean squared errors (MSE) for estimating impulse responses, to impose restrictions on both the short-run dynamics as well as on the adjustment parameters. Ignoring restrictions will most often substantially increase the MSE. Using one set of restrictions is typically notably better than no restrictions, but worse than using both. There is no clear winner between restrictions on the short-run dynamics or on the adjustment parameters when using only one set of restrictions. The size and power properties are improved, but not greatly. Finally, we apply our method to two Swedish macroeconomic data sets and find in some cases vastly different results between the models with or without one or both types of restrictions.

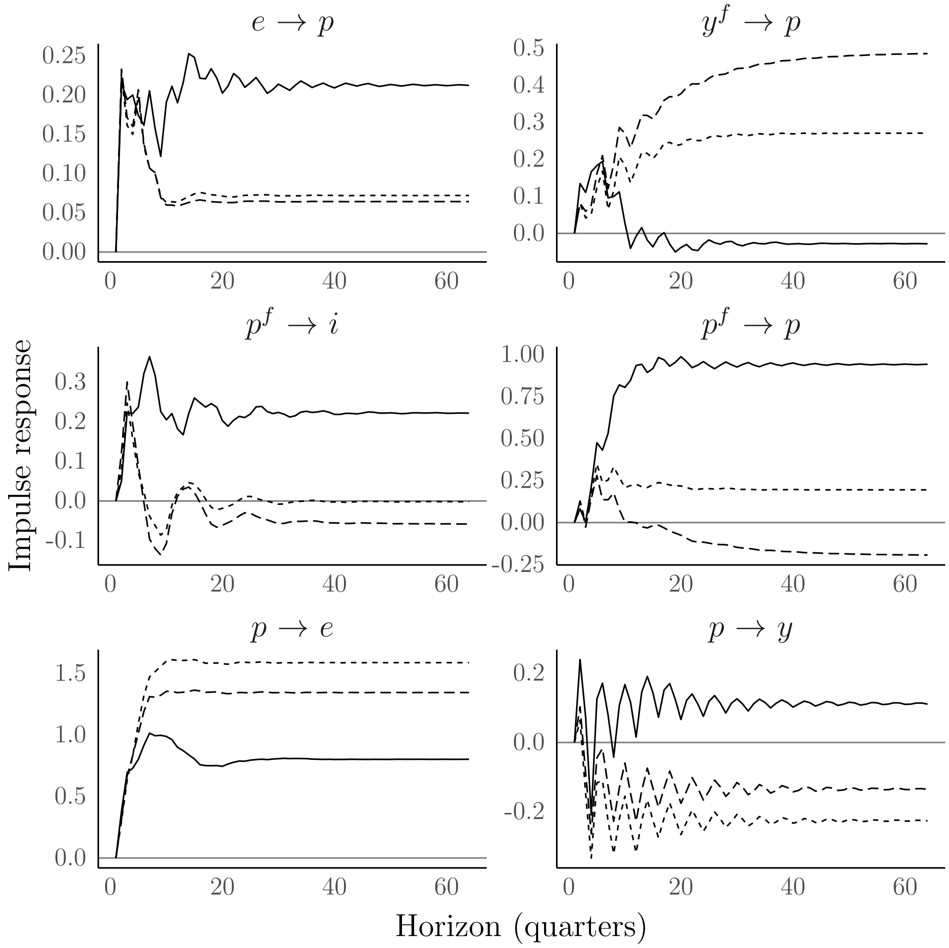

Figure 5: Impulse responses to 1 SD shocks for the Granger non-causal model of the Swedish economy (Section 4.2). Columns denote the origin of the shock and rows the response. The three lines in the figures represent standard VECM (——), VECM with Granger non-causality (−−) and VECM with Granger non-causality and long-run restrictions (- - -).

References

Bjellerup, M., and H. Shahnazarian (2012): "The Interaction between the Financial System and the Real Economy," Discussion paper, Report from the Economic Affairs Department at the Ministry of Finance.

Boswijk, H. (1995): "Identifiability of cointegrated systems," Discussion Paper 78, Tinbergen Institute Discussion Paper.

Boswijk, H. P., and J. A. Doornik (2004): "Identifying, estimating and testing restricted cointegrated systems: An overview," *Statistica Neerlandica*, 58(4), 440–465.

Engle, R. F., and C. W. J. Granger (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55(2), 251–76.

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

Groen, J. J. J., and F. Kleibergen (2003): "Likelihood-based cointegration analysis in panels of vector error-correction models," *Journal of Business & Economic Statistics*, 21(2), 295–318.

Haavelmo, T. (1944): "The Probability Approach in Econometrics," *Econometrica*, 12(Supplement), 1–118.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, NJ.

Jacobs, J. P., and K. F. Wallis (2010): "Cointegration, long-run structural modelling and weak exogeneity: Two models of the UK economy," *Journal of Econometrics*, 158(1), 108 – 116, Twenty Years of Cointegration.

Jacobson, T., P. Jansson, A. Vredin, and A. Warne (2001): "Monetary policy analysis and inflation targeting in a small open economy: a VAR approach*," *Journal of Applied Econometrics*, 16(4), 487–520.

Johansen, S. (1988): "Statistical Analysis of Cointegrated Vectors," *Journal of Economic Dynamics and Control*, 12(2), 231 – 254.

——— (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59(6), 1551–1580.

——— (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

Johansen, S., and K. Juselius (1992): "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK," *Journal of Econometrics*, 53(1), 211 – 244.

——— (1994): "Identification of the long-run and the short-run structure an application to the ISLM model," *Journal of Econometrics*, 63(1), 7 – 36.

Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company.

Lyhagen, J., S. Ekberg, and R. Eidestedt (2015): "Beating the VAR: Improving Swedish GDP Forecasts Using Error and Intercept Corrections," *Journal of Forecasting*, 34(5), 354–363.

MacKinnon, J. G., A. A. Haug, and L. Michelis (1999): "Numerical distribution functions of likelihood ratio tests for cointegration," *Journal of Applied Econometrics*, 14(5), 563–577.

Magnus, J. R. (1978): "Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix," *Journal of Econometrics*, 7(3), 281 – 312.

Magnus, J. R., and H. Neudecker (1979): "The Commutation Matrix: Some Properties and Applications," *Ann. Statist.*, 7(2), 381–394.

Sims, C. (1980): "Macroeconomics and reality," *Econometrica*, 48(1), 1–48.

Toda, H. Y., and P. C. B. Phillips (1993): "Vector Autoregressions and Causality," *Econometrica*, 61(6), 1367–1393.

A Simulation

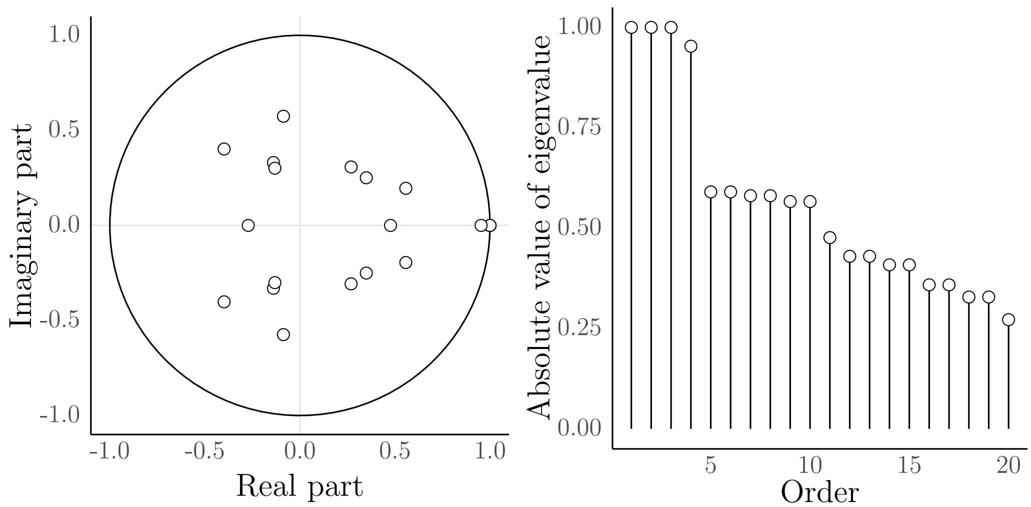

Figure 6: Characteristics of the simulation study's data-generating process. Left: inverse roots of AR characteristic polynomial. Right: absolute values of eigenvalues in decreasing order. As noted in the text, there are two cointegrating relations and thus three unit roots.

B Impulse responses

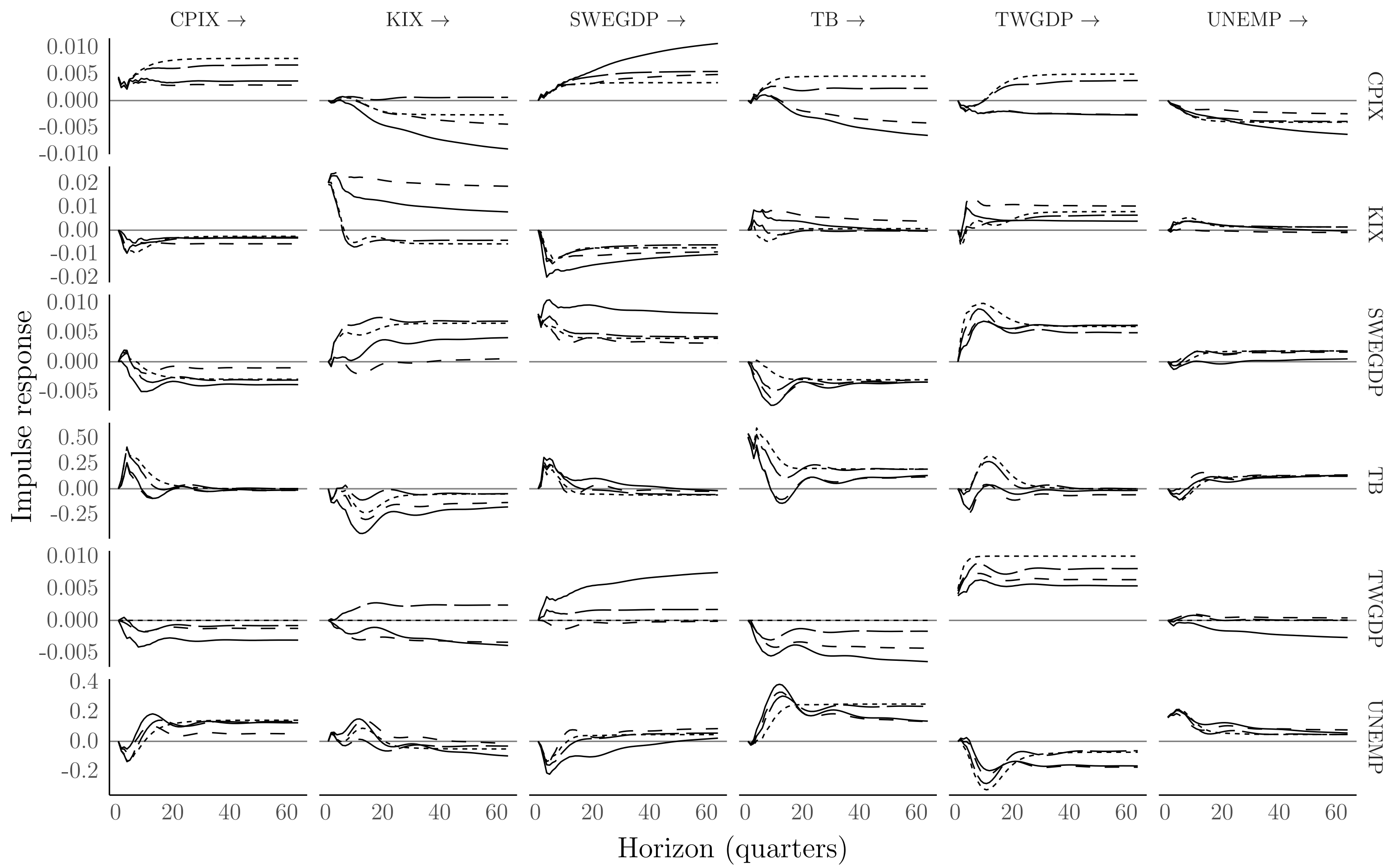

Figure 7: Impulse responses to 1 SD shocks for the weakly exogenous model of the Swedish economy (Section 4.1). Columns denote the origin of the shock and rows the response. The four lines in the figures represent standard VECM (——), VECM with weak exogeneity (– –), VECM with short-run restrictions (- - -) and VECM with short-run restrictions and weak exogeneity (·········).

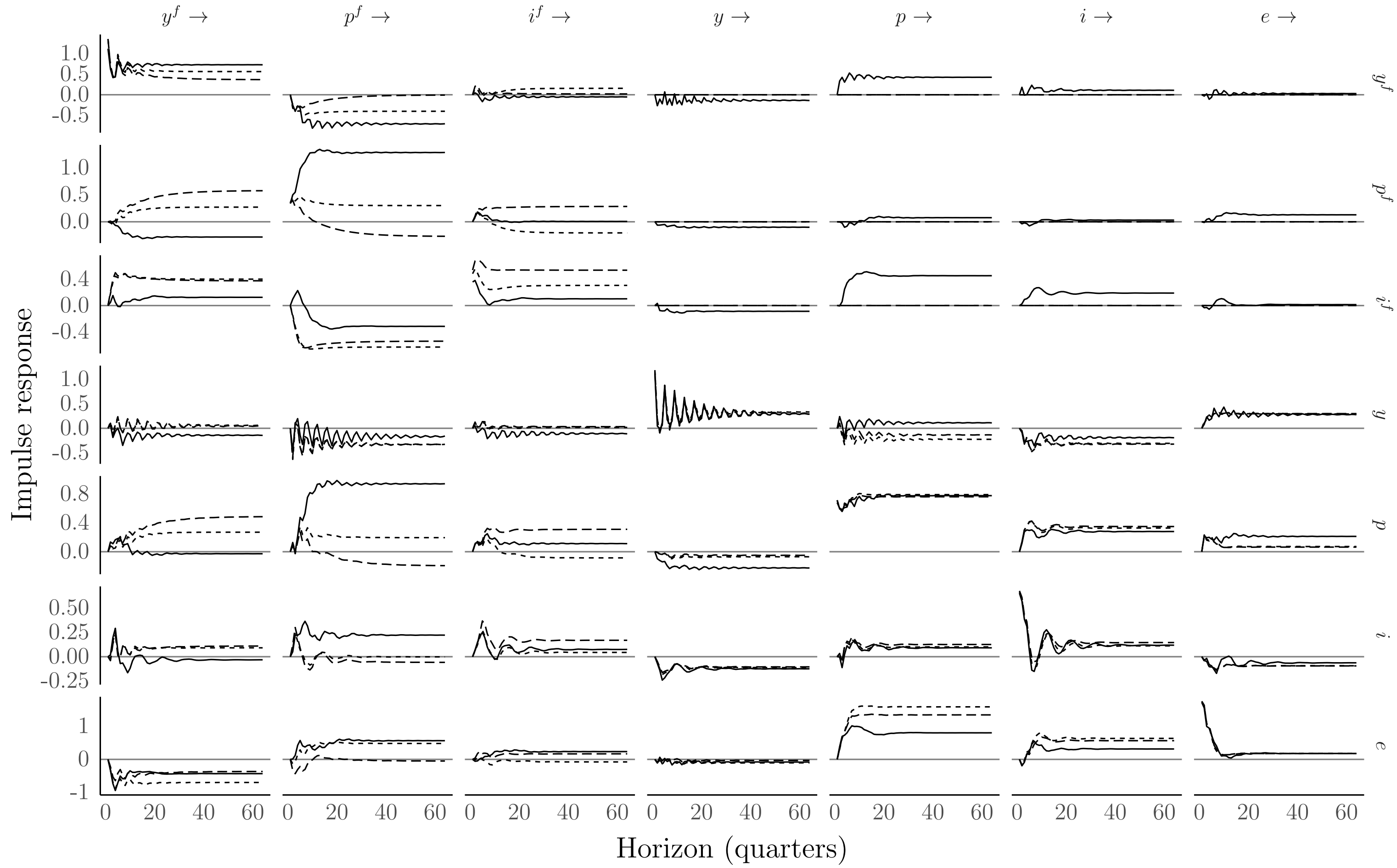

Figure 8: Impulse responses to 1 SD shocks for the Granger non-causal model of the Swedish economy (Section 4.2). Columns denote the origin of the shock and rows the response. The three lines in the figures represent standard VECM (——), VECM with Granger non-causality (– –) and VECM with Granger non-causality and long-run restrictions (- - -).

C Proof

Following Magnus (1978); Groen and Kleibergen (2003) the asymptotic distributions of the seemingly unrelated regressions-type estimators used are the same after one iteration as after full convergence. Moreover, because the model parameters can be consistently estimated in the full model and the iterative procedure produces non-decreasing steps on the log-likelihood surface (Boswijk, 1995; Groen and Kleibergen, 2003; Boswijk and Doornik, 2004) the iterative procedure is also consistent.

To derive the asymptotic distributions, we first introduce a lemma with standard results.

Lemma 1. *Let $W_k(u) = \Omega^{1/2}B_k(u)$ where $B_k(u)$ is a standard k-dimensional Brownian motion. Furthermore, let $C = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)\alpha_\perp'$ where $\Gamma = I - \sum_{i=1}^{p-1}\Gamma_i$ and let $G_k(u) = CW_k(u)$. Moreover, define $\ddot{\Upsilon}_T = \mathrm{diag}(T^{-1}I_k, T^{-1/2}I_{p(k-s)})$. Then*

$$T^{-1/2}\beta_\perp' y_{[Tu]} \xrightarrow{d} CW_k(u) \tag{37}$$

$$T^{-1/2}\Delta y_{[Tu]} \xrightarrow{p} 0 \tag{38}$$

$$\ddot{\Upsilon}_T\tilde{Y}_{-1}'\tilde{Y}_{-1}\ddot{\Upsilon}_T \xrightarrow{p} \begin{pmatrix} \int G_k(u)G_k(u)'du & 0 \\ 0 & \Sigma_{00} \end{pmatrix} \tag{39}$$

$$\ddot{\Upsilon}_T\tilde{Y}_{-1}'\varepsilon \xrightarrow{d} \begin{pmatrix} \int G_k dW_k' \\ \xi \end{pmatrix} \tag{40}$$

$$T^{-1}\beta'\tilde{y}_{-1}'\tilde{y}_{-1}\beta \xrightarrow{p} \Sigma_{\beta\beta} \tag{41}$$

$$T^{-1}\beta'\tilde{y}_{-1}'\Delta\tilde{Y} \xrightarrow{p} \Sigma_{\beta 0} \tag{42}$$

$$T^{-1}\Delta\tilde{Y}'\Delta\tilde{Y} \xrightarrow{p} \Sigma_{00} \tag{43}$$

$$T^{-1/2}\begin{pmatrix} \beta'\tilde{y}_{-1}'\tilde{\varepsilon} \\ \Delta Y'\tilde{\varepsilon} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \zeta \\ \xi \end{pmatrix} \tag{44}$$

where

$$\mathrm{vec}\begin{pmatrix} \zeta \\ \xi \end{pmatrix} \sim N_{k(r+pk)}(0, \Omega \otimes \Sigma) \tag{45}$$

$$\Sigma = \begin{pmatrix} \Sigma_{\beta\beta} & \Sigma_{\beta 0} \\ \Sigma_{0\beta} & \Sigma_{00} \end{pmatrix}. \tag{46}$$

Proof. These results are standard in the literature and follow from Johansen (1995, Lemma 10.2–10.3) and Hamilton (1994, Proposition 18.1). □

Before deriving the asymptotic distributions, a note on notation is in order. We let $(\Omega^{-1})_{ij}$ denote the (i,j)th block of Ω^{-1} where $i,j = 1,2$. Likewise, $(\Omega^{-1})_{i\cdot}$ refers to the ith block of rows across all columns, i.e. $\begin{bmatrix}(\Omega^{-1})_{i1} & (\Omega^{-1})_{i2}\end{bmatrix}$, and vice versa for $(\Omega^{-1})_{\cdot j}$. A similar notation is also used for α, where $\alpha_{\cdot 2}$ refers to the second block of columns (across all rows), as well as for $\Sigma_{\beta\beta}$, where $\Sigma_{\beta\beta,11}$ is the top-left block and $\Sigma_{\beta\beta,1\cdot}$ the upper block of $\Sigma_{\beta\beta}$.

C.1 Asymptotic distribution of $\hat{\pi}_\alpha$ under Granger non-causality

Let

$$\alpha = \left[\begin{array}{c|c} \alpha_{11} & \alpha_{12} = 0_{s\times r_2} \\ \hline \alpha_{21} & \alpha_{22} \end{array}\right] \tag{47}$$

where α_{11} is $s \times r_1$, α_{21} is $(k-s)\times r_1$ and α_{22} is $(k-s)\times r_2$. r_1 denotes cointegrating relations which enter the foreign variables. For convenience, let also $\alpha_1 = (\alpha_{11}, \alpha_{12})$, $\alpha_2 = (\alpha_{21}, \alpha_{22})$ and $\Gamma_d = \begin{bmatrix}\Gamma_{221} & \cdots & \Gamma_{22p}\end{bmatrix}$.

The zero restrictions can be imposed by the following decomposition

$$\operatorname{vec}(\Pi') = \begin{bmatrix} \operatorname{vec}(\alpha\beta') \\ \operatorname{vec}\begin{bmatrix} 0 \\ \Gamma_d \end{bmatrix} \end{bmatrix} = \underbrace{\begin{bmatrix} (\beta \otimes I_k) & 0_{rk\times p(k-s)^2} \\ 0_{pk(k-s)\times rk} & I_{p(k-s)} \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s} \end{bmatrix} \end{bmatrix}}_{F_{1,\beta}(\beta)} \tag{48}$$

$$\times \underbrace{\begin{bmatrix} K_{r,k}\begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} \\ 0_{r_2\times r_1} \end{bmatrix} \\ 0_{r(k-s)\times r_1 s} \end{bmatrix} K_{s,r_1} & I_r \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s} \end{bmatrix} & 0_{rk\times p(k-s)^2} \\ 0_{p(k-s)^2\times r_1 s} & 0_{p(k-s)^2\times r(k-s)} & I_{p(k-s)^2} \end{bmatrix}}_{F_{2,\beta}} \begin{bmatrix} \operatorname{vec}(\alpha_{11}) \\ \operatorname{vec}(\alpha_2) \\ \operatorname{vec}(\Gamma_d) \end{bmatrix}. \tag{49}$$

Let $F_{0,\beta}(\beta) = F_{1,\beta}(\beta)F_{2,\beta}$. Using that

$$\tilde{Y}'_{-1} \otimes I_k = \begin{bmatrix} \tilde{y}'_{-1} \otimes I_k \\ \Delta\tilde{Y}' \otimes I_k \end{bmatrix} \tag{50}$$

and

$$F_{1,\beta}(\beta)'(\tilde{Y}'_{-1} \otimes I_k)K_{T,k} = \begin{bmatrix} \beta'\tilde{y}'_{-1} \otimes I_k \\ \Delta Y' \otimes (0_{k-s\times s}, I_{k-s}) \end{bmatrix} K_{T,k} \tag{51}$$

gives

$$F_{0,\beta}(\beta)'(I_k \otimes \tilde{Y}'_{-1}) = F'_{2,\beta}F_{1,\beta}(\beta)'(\tilde{Y}'_{-1} \otimes I_k)K_{T,k} \tag{52}$$

$$= F'_{2,\beta}\begin{bmatrix} \beta'\tilde{y}'_{-1} \otimes I_k \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{T,k} \tag{53}$$

$$= \begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} & 0_{r_1 s\times r(k-s)} \end{bmatrix} K_{k,r} & 0_{r_1 s\times p(k-s)^2} \\ I_r \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} & 0_{r(k-s)\times p(k-s)^2} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)^2} \end{bmatrix} \begin{bmatrix} \beta'\tilde{y}'_{-1} \otimes I_k \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{T,k} \tag{54}$$

$$= \begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} & 0_{r_1 s\times r(k-s)} \end{bmatrix} K_{k,r}(\beta'\tilde{y}'_{-1} \otimes I_k) \\ \left(I_r \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix}\right) \otimes (\beta'\tilde{y}'_{-1} \otimes I_k) \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{T,k} \tag{55}$$

$$= \begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} & 0_{r_1 s\times r(k-s)} \end{bmatrix} \begin{bmatrix} I_s \otimes \beta'\tilde{y}'_{-1} & 0 \\ 0 & I_{k-s} \otimes \beta'\tilde{y}'_{-1} \end{bmatrix} K_{k,T} \\ \beta'\tilde{y}'_{-1} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{T,k} \tag{56}$$

$$= \begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \beta'_1\tilde{y}'_{-1} & 0_{r_1 s\times T(k-s)} \end{bmatrix} K_{k,T} \\ \beta'\tilde{y}'_{-1} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{T,k} \tag{57}$$

and it follows that

$$T^{-1}F_{0,\beta}(\beta)'\left[\Omega^{-1} \otimes \left(\tilde{Y}'_{-1}\tilde{Y}_{-1}\right)\right]F_{0,\beta}(\beta) \tag{58}$$

$$= T^{-1}F'_{2,\beta}F_{1,\beta}(\beta)'(\tilde{Y}'_{-1} \otimes I_k)K_{T,k}\left(\Omega^{-1} \otimes I_T\right)K_{k,T}(\tilde{Y}_{-1} \otimes I_k)F_{1,\beta}(\beta)F_{2,\beta} \tag{59}$$

$$= T^{-1}\begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \beta'_1\tilde{y}'_{-1} & 0_{r_1 s\times T(k-s)} \end{bmatrix} K_{k,T} \\ \beta'\tilde{y}'_{-1} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} (I_T \otimes \Omega^{-1}) \begin{bmatrix} K_{r_1,s}\begin{bmatrix} I_s \otimes \beta'_1\tilde{y}'_{-1} & 0_{r_1 s\times T(k-s)} \end{bmatrix} K_{k,T} \\ \beta'\tilde{y}'_{-1} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \\ \Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix}'. \tag{60}$$

The resulting matrix is a 3×3 symmetric block matrix. The blocks are in turn:

$$(1,1) = K_{r_1,s}\begin{bmatrix} I_s \otimes \beta_1'\tilde{y}_{-1}' & 0_{r_1 s \times T(k-s)}\end{bmatrix} K_{k,T}(I_T \otimes \Omega^{-1})K_{T,k}\begin{bmatrix} I_s \otimes \tilde{y}_{-1}\beta_1 \\ 0_{T(k-s)\times r_1 s}\end{bmatrix} K_{s,r_1} \tag{61}$$

$$= K_{r_1,s}\begin{bmatrix} I_s \otimes \beta_1'\tilde{y}_{-1}' & 0_{r_1 s \times T(k-s)}\end{bmatrix}\begin{bmatrix} (\Omega^{-1})_{11}\otimes I_T & (\Omega^{-1})_{12}\otimes I_T \\ (\Omega^{-1})_{21}\otimes I_T & (\Omega^{-1})_{22}\otimes I_T\end{bmatrix}\begin{bmatrix} I_s \otimes \tilde{y}_{-1}\beta_1 \\ 0_{T(k-s)\times r_1 s}\end{bmatrix} K_{s,r_1} \tag{62}$$

$$= K_{r_1,s}(I_s \otimes \beta_1'\tilde{y}_{-1}')((\Omega^{-1})_{11}\otimes I_T)(I_s \otimes \tilde{y}_{-1}\beta_1)K_{s,r_1} \tag{63}$$

$$= \beta_1'\tilde{y}_{-1}'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{11} \tag{64}$$

$$(2,1) = \left(\beta'\tilde{y}_{-1}' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix}\right)(I_T \otimes \Omega^{-1})K_{T,k}\begin{bmatrix} I_s \otimes \tilde{y}_{-1}\beta_1 \\ 0_{T(k-s)\times r_1 s}\end{bmatrix} K_{s,r_1} \tag{65}$$

$$= K_{r,k-s}\left(\begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix} \otimes \beta'\tilde{y}_{-1}'\right)\begin{bmatrix} (\Omega^{-1})_{11}\otimes \tilde{y}_{-1}\beta_1 \\ (\Omega^{-1})_{21}\otimes \tilde{y}_{-1}\beta_1\end{bmatrix} K_{s,r_1} \tag{66}$$

$$= \beta'\tilde{y}_{-1}'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{21} \tag{67}$$

$$(3,1) = \left(\Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix}\right)(I_T \otimes \Omega^{-1})K_{T,k}\begin{bmatrix} I_s \otimes \tilde{y}_{-1}\beta_1 \\ 0_{T(k-s)\times r_1 s}\end{bmatrix} K_{s,r_1} \tag{68}$$

$$= K_{r,k-s}\left(\begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix} \otimes \Delta Y'\right)\begin{bmatrix} (\Omega^{-1})_{11}\otimes \tilde{y}_{-1}\beta_1 \\ (\Omega^{-1})_{21}\otimes \tilde{y}_{-1}\beta_1\end{bmatrix} K_{s,r_1} \tag{69}$$

$$= \Delta Y'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{21} \tag{70}$$

$$(2,2) = \left(\beta'\tilde{y}_{-1}' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix}\right)(I_T \otimes \Omega^{-1})\left(\tilde{y}_{-1}\beta \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s}\end{bmatrix}\right) \tag{71}$$

$$= \beta'\tilde{y}_{-1}'\tilde{y}_{-1}\beta \otimes (\Omega^{-1})_{22} \tag{72}$$

$$(3,2) = \left(\Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix}\right)(I_T \otimes \Omega^{-1})\left(\tilde{y}_{-1}\beta \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s}\end{bmatrix}\right) \tag{73}$$

$$= \Delta Y'\tilde{y}_{-1}\beta \otimes (\Omega^{-1})_{22} \tag{74}$$

$$(3,3) = \left(\Delta Y' \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s}\end{bmatrix}\right)(I_T \otimes \Omega^{-1})\left(\Delta Y \otimes \begin{bmatrix} 0_{s\times k-s} \\ I_{k-s}\end{bmatrix}\right) \tag{75}$$

$$= \Delta Y'\Delta Y \otimes (\Omega^{-1})_{22} \tag{76}$$

resulting in

$$T^{-1}F_{0,\beta}(\beta)\left[\Omega^{-1}\otimes\left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]F_{0,\beta}(\beta) \tag{77}$$

$$= T^{-1}\begin{bmatrix} \beta_1'\tilde{y}_{-1}'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{11} & & \\ \beta'\tilde{y}_{-1}'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{21} & \beta'\tilde{y}_{-1}'\tilde{y}_{-1}\beta \otimes (\Omega^{-1})_{22} & \\ \Delta Y'\tilde{y}_{-1}\beta_1 \otimes (\Omega^{-1})_{21} & \Delta Y'\tilde{y}_{-1}\beta \otimes (\Omega^{-1})_{22} & \Delta Y'\Delta Y \otimes (\Omega^{-1})_{22}\end{bmatrix}. \tag{78}$$

By Lemma 1

$$T^{-1}F_{0,\beta}(\beta)\left[\Omega^{-1}\otimes\left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right]F_{0,\beta}(\beta) \xrightarrow{p} \begin{bmatrix} \Sigma_{\beta\beta,11}\otimes(\Omega^{-1})_{11} & \Sigma_{\beta\beta,1\cdot}\otimes(\Omega^{-1})_{12} & \Sigma_{\beta 0,1\cdot}\otimes(\Omega^{-1})_{12} \\ \Sigma_{\beta\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{\beta\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{\beta 0}\otimes(\Omega^{-1})_{22} \\ \Sigma_{0\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{0\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{00}\otimes(\Omega^{-1})_{22}\end{bmatrix}. \tag{79}$$

Furthermore,

$$
\begin{aligned}
& F_{1,\beta}(\beta)' K_{k+p(k-s),k} \left(\Omega^{-1} \otimes \tilde{Y}_{-1}'\right) \operatorname{vec}(\tilde{\varepsilon}) && (80)\\
&= F_{1,\beta}(\beta)' \left(\tilde{Y}_{-1}' \otimes \Omega^{-1}\right) K_{T,k} \operatorname{vec}(\tilde{\varepsilon}) && (81)\\
&= \begin{bmatrix} \beta' \tilde{y}_{-1}' \otimes \Omega^{-1} \\ \left(I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix}\right) \left(\Delta Y' \otimes \Omega^{-1}\right) \end{bmatrix} K_{T,k} \operatorname{vec}(\tilde{\varepsilon}) && (82)\\
&= \begin{bmatrix} I_{rk} & 0 \\ 0 & I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} \left(\begin{bmatrix} \beta' \tilde{y}_{-1}' \\ \Delta Y' \end{bmatrix} \otimes \Omega^{-1} \right) K_{T,k} \operatorname{vec}(\tilde{\varepsilon}) && (83)\\
&= \begin{bmatrix} I_{rk} & 0 \\ 0 & I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{r+p(k-s),k} \left(\Omega^{-1} \otimes \begin{bmatrix} \beta' \tilde{y}_{-1}' \\ \Delta Y' \end{bmatrix} \right) \operatorname{vec}(\tilde{\varepsilon}) && (84)\\
&= \begin{bmatrix} I_{rk} & 0 \\ 0 & I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} K_{r+p(k-s),k} \left(\Omega^{-1} \otimes I_{r+p(k-s)} \right) \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (85)\\
&= \begin{bmatrix} I_{rk} & 0 \\ 0 & I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} \end{bmatrix} \left(I_{r+p(k-s)} \otimes \Omega^{-1} \right) K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (86)\\
&= \begin{bmatrix} I_r \otimes \Omega^{-1} & 0 \\ 0 & \left(I_{p(k-s)} \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix}\right) \left(I_{p(k-s)} \otimes \Omega^{-1}\right) \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (87)\\
&= \begin{bmatrix} I_r \otimes \Omega^{-1} & 0 \\ 0 & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (88)
\end{aligned}
$$

and so

$$
\begin{aligned}
& F_{0,\beta}(\beta)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}' \tilde{\varepsilon}\right) && (89)\\
&= F_{2,\beta}' F_{1,\beta}(\beta)' K_{k+p(k-s),k} \left(\Omega^{-1} \otimes \tilde{Y}_{-1}'\right) \operatorname{vec}(\tilde{\varepsilon}) && (90)\\
&= F_{2,\beta}' \begin{bmatrix} I_r \otimes \Omega^{-1} & 0_{rk\times pk(k-s)} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (91)\\
&= \begin{bmatrix} K_{r_1,s} \begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} & 0_{r_1 s\times r(k-s)} \end{bmatrix} K_{k,r} & 0_{r_1 s\times p(k-s)^2} \\ I_r \otimes \begin{bmatrix} 0_{k-s\times s} & I_{k-s} \end{bmatrix} & 0_{r(k-s)\times p(k-s)^2} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)^2} \end{bmatrix} && (92)\\
&\times \begin{bmatrix} I_r \otimes \Omega^{-1} & 0_{rk\times pk(k-s)} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (93)\\
&= \begin{bmatrix} K_{r_1,s} \begin{bmatrix} I_s \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} & 0_{r_1 s\times r(k-s)} \end{bmatrix} (\Omega^{-1} \otimes I_r) K_{r,k} & 0_{r_1 s\times pk(k-s)} \\ I_r \otimes (\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} && (94)\\
&\times K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix} && (95)\\
&= \begin{bmatrix} K_{r_1,s} \left(\Omega_{1\cdot}^{-1} \otimes \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix}\right) K_{r,k} & 0_{r_1 s\times pk(k-s)} \\ I_r \otimes (\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix}. && (96)\\
&= \begin{bmatrix} \begin{bmatrix} I_{r_1} & 0_{r_1\times r_2} \end{bmatrix} \otimes \Omega_{1\cdot}^{-1} & 0_{r_1 s\times pk(k-s)} \\ I_r \otimes (\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)} \\ 0_{p(k-s)^2\times rk} & I_{p(k-s)} \otimes (\Omega^{-1})_{2\cdot} \end{bmatrix} K_{r+p(k-s),k} \operatorname{vec} \begin{bmatrix} \beta' \tilde{y}_{-1}' \tilde{\varepsilon} \\ \Delta Y' \tilde{\varepsilon} \end{bmatrix}. && (97)
\end{aligned}
$$

Consequently, by Lemma 1

$$T^{-1/2}F_{0,\beta}(\beta)'\left(\Omega^{-1}\otimes I_{k+p(k-s)}\right)\operatorname{vec}\left(\tilde{Y}_{-1}'\tilde{\varepsilon}\right)\xrightarrow{d}$$
$$N_{r_1 s+r(k-s)+p(k-s)^2}\left(0,\begin{bmatrix}\Sigma_{\beta\beta,11}\otimes(\Omega^{-1})_{11} & \Sigma_{\beta\beta,1\cdot}\otimes(\Omega^{-1})_{12} & \Sigma_{\beta 0,1\cdot}\otimes(\Omega^{-1})_{12}\\ \Sigma_{\beta\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{\beta\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{\beta 0}\otimes(\Omega^{-1})_{22}\\ \Sigma_{0\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{0\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{00}\otimes(\Omega^{-1})_{22}\end{bmatrix}\right). \tag{98}$$

The asymptotic covariance matrix in (98) follows from $T^{-1/2}\operatorname{vec}\begin{bmatrix}\beta'\tilde{y}_{-1}'\tilde{\varepsilon}\\ \Delta Y'\tilde{\varepsilon}\end{bmatrix}\xrightarrow{d}\operatorname{vec}\begin{bmatrix}\zeta\\ \xi\end{bmatrix}$ in (97) by Lemma 1, where also the variance of $K_{r+p(k-s),k}\operatorname{vec}\begin{bmatrix}\zeta\\ \xi\end{bmatrix}=\operatorname{vec}\begin{bmatrix}\zeta' & \xi'\end{bmatrix}$ is $\Sigma\otimes\Omega$. The variance in the asymptotic distribution in (98) then comes from

$$\begin{bmatrix}\begin{bmatrix}I_{r_1} & 0_{r_1\times r_2}\end{bmatrix}\otimes\Omega_{1\cdot}^{-1} & 0_{r_1 s\times pk(k-s)}\\ I_r\otimes(\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)}\\ 0_{p(k-s)^2\times rk} & I_{p(k-s)}\otimes(\Omega^{-1})_{2\cdot}\end{bmatrix}\Sigma\otimes\Omega\begin{bmatrix}\begin{bmatrix}I_{r_1} & 0_{r_1\times r_2}\end{bmatrix}\otimes\Omega_{1\cdot}^{-1} & 0_{r_1 s\times pk(k-s)}\\ I_r\otimes(\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)}\\ 0_{p(k-s)^2\times rk} & I_{p(k-s)}\otimes(\Omega^{-1})_{2\cdot}\end{bmatrix}' \tag{99}$$
$$=\begin{bmatrix}\Sigma_{\beta\beta,1\cdot}\otimes\begin{bmatrix}I_s & 0_{s\times k-s}\end{bmatrix} & \Sigma_{\beta 0,1\cdot}\otimes\begin{bmatrix}I_s & 0_{s\times k-s}\end{bmatrix}\\ \Sigma_{\beta\beta}\otimes\begin{bmatrix}0_{k-s\times s} & I_{k-s}\end{bmatrix} & \Sigma_{\beta 0}\otimes\begin{bmatrix}0_{k-s\times s} & I_{k-s}\end{bmatrix}\\ \Sigma_{0\beta}\otimes\begin{bmatrix}0_{k-s\times s} & I_{k-s}\end{bmatrix} & \Sigma_{00}\otimes\begin{bmatrix}0_{k-s\times s} & I_{k-s}\end{bmatrix}\end{bmatrix}\begin{bmatrix}\begin{bmatrix}I_{r_1} & 0_{r_1\times r_2}\end{bmatrix}\otimes\Omega_{1\cdot}^{-1} & 0_{r_1 s\times pk(k-s)}\\ I_r\otimes(\Omega^{-1})_{2\cdot} & 0_{r(k-s)\times pk(k-s)}\\ 0_{p(k-s)^2\times rk} & I_{p(k-s)}\otimes(\Omega^{-1})_{2\cdot}\end{bmatrix}' \tag{100}$$
$$=\begin{bmatrix}\Sigma_{\beta\beta,11}\otimes(\Omega^{-1})_{11} & \Sigma_{\beta\beta,1\cdot}\otimes(\Omega^{-1})_{12} & \Sigma_{\beta 0,1\cdot}\otimes(\Omega^{-1})_{12}\\ \Sigma_{\beta\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{\beta\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{\beta 0}\otimes(\Omega^{-1})_{22}\\ \Sigma_{0\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{0\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{00}\otimes(\Omega^{-1})_{22}\end{bmatrix}. \tag{101}$$

As the (scaled) estimator can be written as the inverse of (79) times (98), the asymptotic distribution of the estimator is

$$\sqrt{T}(\hat{\pi}_\alpha-\pi_\alpha)\xrightarrow{d}N_{r_1 s+r(k-s)+p(k-s)^2}\left(0,\begin{bmatrix}\Sigma_{\beta\beta,11}\otimes(\Omega^{-1})_{11} & \Sigma_{\beta\beta,1\cdot}\otimes(\Omega^{-1})_{12} & \Sigma_{\beta 0,1\cdot}\otimes(\Omega^{-1})_{12}\\ \Sigma_{\beta\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{\beta\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{\beta 0}\otimes(\Omega^{-1})_{22}\\ \Sigma_{0\beta,\cdot 1}\otimes(\Omega^{-1})_{21} & \Sigma_{0\beta}\otimes(\Omega^{-1})_{22} & \Sigma_{00}\otimes(\Omega^{-1})_{22}\end{bmatrix}^{-1}\right). \tag{102}$$

C.2 Asymptotic distribution of $\hat{\pi}_\beta$ under Granger non-causality

Let

$$\beta=\left[\begin{array}{c|c}\beta_{11} & \beta_{12}\\ \hline \beta_{21}=0_{k-s\times r_1} & \beta_{22}\end{array}\right] \tag{103}$$

where β_{11} is $s\times r_1$, β_{21} is $k-s\times r_1$, β_{12} is $s\times r_2$ and β_{22} is $(k-s)\times r_2$. r_1 denotes cointegrating relations which enter the foreign variables. For convenience, let also $\beta_1=(\beta_{11},\beta_{12})$ and $\beta_2=(\beta_{21},\beta_{22})$.

The zero restrictions can be enforced by the following decomposition

$$\operatorname{vec}(\Pi)'=\begin{bmatrix}\operatorname{vec}(\alpha\beta')\\ \operatorname{vec}\begin{bmatrix}0\\ \Gamma_d\end{bmatrix}\end{bmatrix} \tag{104}$$
$$=\underbrace{\left[\begin{array}{c|c|c}\begin{bmatrix}I_s\\ 0_{k-s\times s}\end{bmatrix}\otimes\alpha & \begin{matrix}0_{ks\times r_2(k-s)}\\ I_{k-s}\otimes\alpha\begin{bmatrix}0_{r_1\times r_2}\\ I_{r_2}\end{bmatrix}\end{matrix} & 0_{r(k-s)\times p(k-s)^2}\\ \hline 0_{pk(k-s)\times rs} & 0_{pk(k-s)\times r_2(k-s)} & I_{p(k-s)}\otimes\begin{bmatrix}0_{s\times k-s}\\ I_{k-s}\end{bmatrix}\end{array}\right]}_{F_{2,\alpha}(\alpha)}\begin{bmatrix}\operatorname{vec}(\beta_1')\\ \operatorname{vec}(\beta_{22}')\\ \operatorname{vec}(\Gamma_d)\end{bmatrix}. \tag{105}$$

Let

$$F_{0,\alpha}(\alpha) = \left(I \otimes \tilde{Y}_{-1}'\tilde{Y}_{-1}\right) F_{1,\alpha}(\alpha) = \left(I \otimes \tilde{Y}_{-1}'\tilde{Y}_{-1}\right) K_{k,k+p(k-s)} F_{2,\alpha}(\alpha). \tag{106}$$

It is then possible to rewrite (21) as

$$\hat{\pi}_\beta = \pi_\beta + \left\{F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right] F_{1,\alpha}(\alpha)\right\}^{-1} F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\tilde{\varepsilon}\right) \tag{107}$$

by using the identity $(I \otimes A)(B \otimes A^{-1})(I \otimes A) = (B \otimes A)$. Consider now $\Upsilon_T = \operatorname{diag}(T^{-1} I_{rs+r_2(k-s)}, T^{-1/2} I_{p(k-s)^2})$ and

$$\Upsilon_T F_{1,\alpha}(\alpha)' = \Upsilon_T F_{2,\alpha}(\alpha)' K_{k+p(k-s),k} = F_{2,\alpha}(\alpha)'(\ddot{\Upsilon}_T \otimes I_k) K_{k+p(k-s),k}. \tag{108}$$

Additionally, we can write

$$(\Omega^{-1} \otimes \tilde{Y}_{-1}'\tilde{Y}) = (I_k \otimes \tilde{Y}_{-1}')(\Omega^{-1} \otimes I_T)(I_k \otimes \tilde{Y}_{-1}). \tag{109}$$

By Magnus and Neudecker (1979, Theorem 3.1(viii)) we also have

$$K_{k+p(k-s),k}(I_k \otimes \tilde{Y}_{-1}') = (\tilde{Y}_{-1}' \otimes I_k) K_{T,k} \tag{110}$$

and hence it follows that

$$\begin{aligned} (\ddot{\Upsilon}_T \otimes I_k) K_{k+p(k-s),k}(I_k \otimes \tilde{Y}_{-1}') &= (\ddot{\Upsilon}_T \otimes I_k)(\tilde{Y}_{-1}' \otimes I_k) K_{T,k} \\ &= (\ddot{\Upsilon}_T \tilde{Y}_{-1}' \otimes I_k) K_{T,k} \end{aligned} \tag{111}$$

By symmetry, (108) and (111) imply that

$$\left\{\Upsilon_T F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right] F_{1,\alpha}(\alpha) \Upsilon_T\right\}^{-1} \tag{112}$$

$$= \left\{F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \left(\ddot{\Upsilon}_T \tilde{Y}_{-1}'\tilde{Y}_{-1} \ddot{\Upsilon}_T\right)\right] F_{1,\alpha}(\alpha)\right\}^{-1} \tag{113}$$

By Lemma 1 and the continuous mapping theorem, we obtain

$$\begin{aligned} &\left\{\Upsilon_T F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right] F_{1,\alpha}(\alpha) \Upsilon_T\right\}^{-1} \\ \xrightarrow{p} &\left\{F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \begin{pmatrix} \int G_k(u) G_k(u)' du & 0_{k \times p(k-s)} \\ 0_{p(k-s) \times k} & \Sigma_{00} \end{pmatrix}\right] F_{1,\alpha}(\alpha)\right\}^{-1} \end{aligned} \tag{114}$$

Furthermore, we can write

$$\begin{aligned} &\Upsilon_T F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right) \\ = &F_{2,\alpha}(\alpha)'(\ddot{\Upsilon}_T \otimes I_k) K_{k+p(k-s),k} \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right), \end{aligned} \tag{115}$$

where also

$$\begin{aligned} &(\ddot{\Upsilon}_T \otimes I_k) K_{k+p(k-s),k} \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right) \\ &= K_{k+p(k-s),k}(I_k \otimes \ddot{\Upsilon}_T) \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right) \\ &= K_{k+p(k-s),k}(\Omega^{-1} \otimes \ddot{\Upsilon}_T) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right) \\ &= K_{k+p(k-s),k} \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\ddot{\Upsilon}_T \tilde{Y}_{-1}'\varepsilon\right). \end{aligned} \tag{116}$$

Thus, we have by (115), (116) and the continuous mapping theorem

$$\begin{aligned}&\Upsilon_T F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}' \varepsilon\right) \\ &\xrightarrow{d} F_{1,\alpha}(\alpha) \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\begin{pmatrix} \int G_k dW_k' \\ \xi \end{pmatrix}.\end{aligned} \tag{117}$$

Taking (114) and (117) together results in

$$\begin{aligned}\Upsilon_T^{-1}\left(\hat{\pi}_\beta - \pi_\beta\right) &= \left\{\Upsilon_T F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \left(\tilde{Y}_{-1}'\tilde{Y}_{-1}\right)\right] F_{1,\alpha}(\alpha)\Upsilon_T\right\}^{-1} \\ &\quad\times \Upsilon_T F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\left(\tilde{Y}_{-1}'\varepsilon\right) \\ &\xrightarrow{d} \left\{F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \begin{pmatrix} \int G_k(u)G_k(u)'du & 0_{k\times p(k-s)} \\ 0_{p(k-s)\times k} & \Sigma \end{pmatrix}\right] F_{1,\alpha}(\alpha)\right\}^{-1} \\ &\quad\times F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\begin{pmatrix} \int G_k dW_k' \\ \xi \end{pmatrix}\end{aligned} \tag{118}$$

Next,

$$\begin{aligned}&F_{1,\alpha}(\alpha)' \left[\Omega^{-1} \otimes \begin{pmatrix} \int G_k(u)G_k(u)'du & 0_{k\times p(k-s)} \\ 0_{p(k-s)\times k} & \Sigma_{00} \end{pmatrix}\right] F_{1,\alpha}(\alpha) \\ &= F_{2,\alpha}(\alpha)' \left(\begin{bmatrix} \int G_k(u)G_k(u)'du & 0_{k\times p(k-s)} \\ 0_{p(k-s)\times k} & \Sigma_{00} \end{bmatrix} \otimes \Omega^{-1}\right) F_{2,\alpha}(\alpha) \\ &= F_{2,\alpha}(\alpha)' \begin{bmatrix} \left(\int G_k(u)G_k(u)'du\right) \otimes \Omega^{-1} & 0_{k^2\times pk(k-s)} \\ 0_{pk(k-s)\times k^2} & \Sigma_{00} \otimes \Omega^{-1} \end{bmatrix} F_{2,\alpha}(\alpha) \\ &= \begin{bmatrix} \int G_{k,1}(u)G_{k,1}(u)'du \otimes \alpha'\Omega^{-1}\alpha & \int G_{k,1}(u)G_{k,2}(u)'du \otimes \alpha'\Omega^{-1}\alpha_{\cdot 2} & 0_{rs\times pk(k-s)} \\ \int G_{k,2}(u)G_{k,1}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha & \int G_{k,2}(u)G_{k,2}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha_{\cdot 2} & 0_{rs\times pk(k-s)} \\ 0_{pk(k-s)\times rs} & 0_{pk(k-s)\times r(k-s)} & \Sigma_{00} \otimes (\Omega^{-1})_{22} \end{bmatrix}\end{aligned} \tag{119}$$

Similarly,

$$F_{1,\alpha}(\alpha)' \left(\Omega^{-1} \otimes I_{k+p(k-s)}\right) \operatorname{vec}\begin{bmatrix} \int G_k dW_k' \\ \xi \end{bmatrix} \tag{120}$$

$$= F_{2,\alpha}(\alpha)' \operatorname{vec}\left[\Omega^{-1/2} \int dB_k G_k' \quad \Omega^{-1}\xi'\right] \tag{121}$$

$$= \begin{bmatrix} \left[I_s \quad 0_{s\times k-s}\right] \otimes \alpha' \operatorname{vec}\left(\Omega^{-1/2} \int dB_k G_k'\right) \\ \left[0_{r_2(k-s)\times ks} \quad I_{k-s} \otimes \alpha_{\cdot 2}'\right] \operatorname{vec}\left(\Omega^{-1/2} \int dB_k G_k'\right) \\ I_{p(k-s)} \otimes \left[0_{k-s\times s} \quad I_{k-s}\right] \operatorname{vec}\left(\Omega^{-1}\xi'\right) \end{bmatrix} \tag{122}$$

$$= \begin{bmatrix} \operatorname{vec}\left(\alpha'\Omega^{-1/2} \int dB_k G_{k,1}'\right) \\ \operatorname{vec}\left(\alpha_{\cdot 2}'\Omega^{-1/2} \int dB_k G_{k,2}'\right) \\ \operatorname{vec}\left((\Omega^{-1})_{2\cdot}\xi'\right) \end{bmatrix}. \tag{123}$$

Thus,

$$\Upsilon_T^{-1}(\hat{\pi}_\beta - \pi_\beta) \xrightarrow{d} \tag{124}$$

$$\begin{bmatrix} \int G_{k,1}(u)G_{k,1}(u)'du \otimes \alpha'\Omega^{-1}\alpha & \int G_{k,1}(u)G_{k,2}(u)'du \otimes \alpha'\Omega^{-1}\alpha_{\cdot 2} & 0_{rs\times pk(k-s)} \\ \int G_{k,2}(u)G_{k,1}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha & \int G_{k,2}(u)G_{k,2}(u)'du \otimes \alpha_{\cdot 2}'\Omega^{-1}\alpha_{\cdot 2} & 0_{rs\times pk(k-s)} \\ 0_{pk(k-s)\times rs} & 0_{pk(k-s)\times r(k-s)} & \Sigma_{00} \otimes (\Omega^{-1})_{22} \end{bmatrix}^{-1} \tag{125}$$

$$\times \begin{bmatrix} \operatorname{vec}\left(\alpha'\Omega^{-1/2} \int dB_k G_{k,1}'\right) \\ \operatorname{vec}\left(\alpha_{\cdot 2}'\Omega^{-1/2} \int dB_k G_{k,2}'\right) \\ \operatorname{vec}\left((\Omega^{-1})_{2\cdot}\xi'\right) \end{bmatrix} \tag{126}$$