

Schreiber, Sven

**Conference Paper**

## The estimation uncertainty of permanent-transitory decompositions in co-integrated systems

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik  
- Session: Time Series Analysis, No. C16-V3

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Schreiber, Sven (2014) : The estimation uncertainty of permanent-transitory decompositions in co-integrated systems, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Time Series Analysis, No. C16-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100582>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The estimation uncertainty of permanent-transitory decompositions in co-integrated systems\*

by Sven Schreiber<sup>†</sup>

this version: February 2014

## Abstract

The topic of this paper is the estimation uncertainty of the Stock-Watson and Gonzalo-Granger permanent-transitory decompositions in the framework of the co-integrated vector autoregression. We suggest an approach to construct the confidence interval of the transitory component estimate in a given period (e.g. the latest observation) by conditioning on the observed data in that period. To calculate asymptotically valid confidence intervals we use the delta method and two bootstrap variants. As an illustration we analyze the uncertainty of (US) output gap estimates in a system of output, consumption, and investment.

**Keywords:** transitory components, VECM, delta method, bootstrap

**JEL codes:** C32 (multiple time series), C15 (simulation methods), E32 (business cycles)

---

\*Thanks to participants at the Econometric Society European meeting and EUI Florence for helpful comments, especially Jörg Breitung, Neil Ericsson, Helmut Lütkepohl, and Søren Johansen, as well as to anonymous referees for suggestions which have clearly improved the paper. The author acknowledges the hospitality of the CES at University Paris 1 Panthéon-Sorbonne CNRS, where parts of the early research for this paper were carried out. The numeric calculations have been produced with gretl 1.9 (Cottrell and Lucchetti, 2010), the code will be available as a gretl function package, on request, or from the web page of the author ([www.sven.schreiber.name](http://www.sven.schreiber.name)).

<sup>†</sup>Macroeconomic Policy Institute (IMK) at Hans Boeckler Foundation, and Free University Berlin. E-mail: [mail@sven.schreiber.name](mailto:mail@sven.schreiber.name), fax: +49(0)211-7778-4596, address: Hans-Böckler-Stiftung/IMK, Hans-Böckler-Str. 39, D-40476 Düsseldorf, Germany.

# 1 Introduction

In this paper we suggest an approach to assess the estimation uncertainty of two permanent-transitory (PT) decompositions estimated in a co-integrated VAR framework, namely of the [Stock and Watson \(1988, SW\)](#) and [Gonzalo and Granger \(1995, GG\)](#) methods.

There are many ways to decompose integrated multivariate time series into their unobserved permanent and transitory components, even if we restrict our attention to additive decompositions  $y_t = y_t^{perma} + y_t^{trans}$  (where  $y_t$  is an  $n$ -dimensional time series), and also ruling out univariate decompositions applied separately to the elements of  $y_t$ . The most widespread methods are state-space based unobserved components models (also known as structural time series models, see [Harvey and Shephard, 1993](#), and [Harvey and Proietti, 2005](#)), and decompositions based on co-integrated VARs on the other hand (vector error-correction models, VECM). The leading examples of the VECM based decompositions are the extraction of SW common trends and GG common factors with their corresponding transitory components.

The state-space approach is a powerful and flexible tool which also has the advantage that the Kalman filter provides a way to assess the uncertainty surrounding the estimates of the (smoothed) states and therefore of the permanent component. In contrast, for VECM based measures there has not existed a way to quantify the estimation uncertainty for a given period of interest. It has only been possible so far to assess the significance of the coefficients of the transitory decomposition in general for the whole sample (on average), by standard test procedures. Therefore the goal of this paper is to provide additional tools to quantify the uncertainty around the SW and GG decompositions. The general idea is to condition on the data constellation prevailing in the period of interest and to focus on the parameter estimation uncertainty arising from the remainder of the estimation sample.

These methods may also prove useful since the SW and GG measures only rely

on the available estimated quantities from the VECM via closed-form algebraic expressions, and thus they do not suffer from typical practical problems of iterative numerical methods such as those that have to be used for the estimation of state-space models.

The entire analysis will be conducted conditional on a fixed co-integration rank, as well as treating the lag order as given. This means that the model selection uncertainty will not be captured by our confidence bands, only the parameter uncertainty. While the true co-integration rank may often be known (or at least imposed) a priori due to theoretical restrictions, this knowledge will typically not exist for the true lag order. But such a conditional analysis is a standard approach to construct standard errors for VECM coefficients (including the implied impulse-response coefficients).

After formally introducing the model framework and fixing the notation in the following section, we analyze the uncertainty of the GG and SW transitory components in section 3. There we state the decompositions in a way that is especially useful for our purposes, we explain our conditioning approach, and we derive the covariance matrices of the transitory components by applying the delta method. Afterwards in section 4 we present some bootstrapping variants as alternative ways to assess the estimation uncertainty. In section 5 both approaches are applied to a three-variable dataset inspired by the influential work of [King, Plosser, Stock, and Watson \(1991\)](#), but updated to include more recent data. Section 6 summarizes.

## 2 Framework and assumptions

Consider a standard  $n$ -dimensional VAR with  $p$  lags:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \mu + \varepsilon_t, \quad t = p, \dots, T \quad (1)$$

where the innovations are white noise with covariance matrix  $\Theta$ . We can re-parameterize this system as a VECM:

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{p-1} B_i \Delta y_{t-i} + \mu + \varepsilon_t \quad (2)$$

When co-integration is present, the long-run matrix  $\alpha \beta' = -I + \sum_{i=1}^p A_i$  has reduced rank  $r$  which is the number of linearly independent co-integration relationships (and is also the column rank of the  $n \times r$  matrices  $\alpha$  and  $\beta$ ). The coefficients of the lagged differences are given by  $B_j = -\sum_{i=j+1}^p A_i$ . We define the lag polynomial  $B(L) = I - \sum_{i=1}^{p-1} B_i L^i$ . Because it will be repeatedly needed below, we introduce an abbreviation for the following term:  $Q \equiv B(1) - \alpha \beta'$ .

It is well known that the constant term  $\mu$  can serve two purposes: if unrestricted, it may represent a linear drift term in the levels of the variables, as well as balancing the mean of the co-integrating relations. But if it is restricted as  $\mu = \alpha \mu_0$ , the levels of the data are assumed to be free of linear trend components. In the following, we will deal with the more general case of an unrestricted constant, which is much more popular in economics given the trending behavior of many variables in growing economies. As a further deterministic component it would also be possible for our analysis to allow a linear trend term in the co-integrating relations, because the convergence rate of its estimator is also greater than  $\sqrt{T}$ . (It may be advisable in practical work to normalize the trend term to have mean zero.) Our explicit formulation in this paper focuses on the presented case, however.

Apart from standard regularity conditions like a well-behaved distribution of the residuals  $\varepsilon_t$ , that the process was started in the distant past such that the influence of the initial conditions vanishes, and that standard asymptotic results for the VECM apply, we make the following assumptions:

**Assumption 1.** *All variables are individually  $I(0)$  or  $I(1)$ .*

*This assumption rules out higher integration orders, and also explosive roots.*

**Assumption 2.** *The matrices  $Q$  and  $\beta' Q^{-1} \alpha$  are non-singular (with rank  $n$  and  $r$ , respectively).*

These conditions are also used by [Proietti \(1997\)](#) on whose framework our analysis is based, and they are sufficient to guarantee stability of the state-space representation of the VECM. While the conditions are not strictly necessary for stability,<sup>1</sup> they facilitate our analytical treatment in this paper.

**Assumption 3.** Fixed co-integration rank  $r$ ,  $n > r > 0$ , and fixed lag order  $p$ .

The correct specification of the model is essentially treated as a separate pre-test problem outside of the estimation problem of the VECM and the transitory components of the data. Of course, the true co-integration rank may be known in some contexts.

**Assumption 4.** The co-integration coefficients  $\beta$  are estimated by maximum likelihood (“Johansen procedure”) and are properly normalized and identified.

This assumption serves to achieve super-consistency of the estimates of the co-integration coefficients, see [Paruolo \(1997\)](#). Inter alia it means that identification is achieved by imposing restrictions on  $\beta$ , not on  $\alpha$ , and that no coefficients with a true value of zero are “normalized” to a non-zero value. Other super-consistent estimation methods may be used as well.

In order to state the distribution of the underlying short-run coefficients, we collect the parameters in one matrix  $(\alpha, B_1, \dots, B_{p-1}, \mu)$  and stack the coefficients in the vector  $k = \text{vec}(\alpha, B_1, \dots, B_{p-1}, \mu)$ ; this vector has  $nr + n^2(p-1) + n$  elements that are freely varying.<sup>2</sup> Note that  $\beta$  is not included here because its estimate will be treated as asymptotically fixed given its higher convergence rate ( $T$  instead of

---

<sup>1</sup>I am grateful to an anonymous referee of an earlier version for pointing out this fact. Stability of Proietti’s state-space representation requires that  $I - \mathbf{T}$  is non-singular, where  $\mathbf{T}$  is the transition matrix of the state transition equation. The invertibility of  $Q$  makes it possible to derive  $(I - \mathbf{T})^{-1}$  by using formulas for partitioned matrices, but it is possible to construct cases where  $I - \mathbf{T}$  has full rank despite  $Q$  being singular; in these cases the conditions of the Granger representation theorem are still fulfilled. Therefore the statement by [Hecq, Palm, and Urbain \(2000, p. 516\)](#) that assumption 2 should follow from assumption 1 is slightly misleading. (Of course we have adapted these conditions to match our different notation.)

<sup>2</sup>The only qualification here is given by the standard assumptions that were made about the model class, i.e. the co-integration rank must be preserved and the system must not become integrated of higher order. These requirements are fulfilled in the neighborhood of the true parameters.

$\sqrt{T}$ ), i.e. its variation is asymptotically dominated by the variation of the estimators of the elements of  $k$ . The OLS estimate of this vector  $k$  is asymptotically normally distributed and has  $\sqrt{T}$ -convergence.<sup>3</sup>

*Remark 1.* Standard asymptotics of the underlying coefficients:

$$\sqrt{T}(\hat{k} - k) \xrightarrow{d} N(0, \Omega),$$

where  $\xrightarrow{d}$  denotes convergence in distribution. The covariance matrix  $\Omega$  can be easily estimated as  $\hat{\Theta} \otimes (X'X)^{-1}$  within the standard system OLS estimation once the super-consistent estimate  $\hat{\beta}$  has been determined, where the symbol  $\otimes$  denotes the Kronecker product, and  $X$  is the data matrix of all regressors (including the error correction terms).

*Remark 2.* The influence of the data in any finite number of observation periods on the estimates vanishes asymptotically. Therefore the limit distribution of the estimates  $\hat{k}$  is identical whether or not some observations are removed from the sample.

### 3 Constructing the variances of the transitory components

In this section we will derive the GG and SW decompositions in a representation that is especially suitable for our purposes, and we explain how to apply the classic delta method to the respective transitory components.

As briefly mentioned in the introduction, the idea in this paper is to condition on the observed data relevant for the decomposition in period  $\tau$ . It will become clear that for the GG decomposition this concerns the data  $y_\tau$ , and for the SW decomposition the  $p$  vectors  $y_\tau, \dots, y_{\tau-p+1}$  are involved. We could implement this idea by

---

<sup>3</sup>This also applies to the  $\alpha_\perp$ -directions of the constant term, see [Paruolo \(1997\)](#).

actually removing the conditioning data from the likelihood function; this could be achieved either by using impulse dummies for the corresponding observations, or in the often interesting case of the end of the sample, by simply shortening the sample. However, here we use an approach which makes it unnecessary to re-estimate the model: Taking into account remark 2 we rely on the fact that the conditioning data are negligible relative to the rest of the large sample. Hence, given that in any case our calculated confidence bands are only valid asymptotically, we estimate the model once over the entire sample including the data observations on which we will later condition. This can be regarded as a computational shortcut.

### 3.1 Definition and representation of the GG decomposition

As shown by [Gonzalo and Granger \(1995\)](#), when the permanent and transitory components are assumed to be linear combinations of the contemporaneous values  $y_t$  only, the PT decomposition is uniquely given as follows:

$$y_t = \beta_{\perp} (\alpha'_{\perp} \beta_{\perp})^{-1} \alpha'_{\perp} y_t + \alpha (\beta' \alpha)^{-1} \beta' y_t, \quad (3)$$

where the first part is the non-stationary permanent component, and the second part is the transitory component given by a linear combination of the co-integrating relationships.

We will use the alternative formulation by [Hecq, Palm, and Urbain \(2000\)](#) (based in turn on [Proietti, 1997](#)) with only slightly different notation, which proves especially useful with the Stock-Watson decomposition below. An important projection matrix is given by

$$P = Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} \beta' \quad (4)$$

Since  $\psi_{1t} = P y_t$  is a linear combination of the co-integrating relations  $\beta' y_t$  it is obviously stationary, and it is actually shown by [Proietti \(1997\)](#) that this is just the



GG transitory component:

$$y_t^{transGG} = \psi_{1t} = Py_t \quad (5)$$

This transitory component will in general have a non-zero mean, however. For an economic interpretation it is especially useful to consider a transformation of the transitory component which will have an unconditional expectation of zero, because the sign of that transformed component automatically tells us whether the observed level of a variable is below or above its permanent component. For example the sign of an output gap estimate is important for identifying a recessionary or overheating economy.

To this end we use the expression (again adapted from [Proietti, 1997](#)) for the mean of the co-integrating relationships:

$$E(\beta' y_t) = -(\beta' Q^{-1} \alpha)^{-1} \beta' Q^{-1} \mu, \quad (6)$$

which enables us to calculate the de-meanned transitory component:

$$\begin{aligned} \tilde{\psi}_{1t} &= \psi_{1t} - E(\psi_{1t}) \\ &= Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} (\beta' y_t + [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1} \mu) \\ &= Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} (\beta', [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1} \mu) (y'_t, 1)' \\ &\equiv G(y'_t, 1)' \end{aligned} \quad (7)$$

Of course it is well known how to test the hypothesis that the GG transitory component of a certain variable vanishes completely. From the definitions of the GG decomposition it is clear that this involves a test that the  $i$ -th row of  $\alpha$  is zero, which is a standard test problem given the co-integration rank and the estimated co-integration coefficients. This paper is instead concerned with the uncertainty of

the transitory component at a certain period, assuming that it exists at all.

### 3.2 The Delta method for the GG decomposition

We can express the de-meaned transitory GG component  $\tilde{\psi}_{1\tau}$  in period  $\tau \in \{p, \dots, T\}$  as a function of the underlying short-run coefficient vector  $k$  (whose estimates are  $\sqrt{T}$ -consistent), of the super-consistent co-integration coefficients  $\beta$ , and of the data; since the Gonzalo-Granger transitory component  $\tilde{\psi}_{1\tau}$  only depends on the contemporaneous observations, we only need to condition on  $y_\tau$ :

$$\tilde{\psi}_{1\tau} = f_{GG}(k; \beta, y_\tau) \quad (8)$$

By (7) we have  $f_{GG} = G(y'_t, 1)'$ . Let  $J_{GG} = \partial \tilde{\psi}_{1\tau} / \partial k'$  be the Jacobian matrix of that function with respect to  $k$ , treating the co-integration coefficients  $\hat{\beta}$  as (asymptotically) fixed and conditioning on the data in period  $\tau$ . In the appendix we calculate this Jacobian analytically.<sup>4</sup> With this definition we can state the first result with respect to the estimation uncertainty of the GG transitory component.

**Proposition 1.** *The conditional asymptotic distribution of the GG transitory component estimator for a fixed  $y_\tau$  is given by:*

$$\sqrt{T}(\hat{\psi}_{1\tau} - \tilde{\psi}_{1\tau}) \xrightarrow{d} N(0, J_{GG} \Omega J'_{GG}), \quad (9)$$

The proposition follows directly as an application of the standard delta method. Given the  $T$ -convergence of the co-integration coefficient estimates  $\hat{\beta}$ , their variation is asymptotically dominated by that of the other coefficients and thus formally negligible. The influence of  $y_\tau$  on the estimates is either non-existent (if a dummy variable for period  $\tau$  was used) or asymptotically negligible. A standard system

---

<sup>4</sup>In earlier versions we used gretl's fdjac() function as a numerical approximation. It turned out that the approximation is almost perfect for the Gonzalo-Granger case, and still good in the Stock-Watson case below, where it yielded results that were somewhat more volatile.

OLS estimate  $\hat{\Omega}$  (for a given  $\hat{\beta}$ ) and an estimate of  $J_{GG}$  based on  $\hat{k}$  can be used for a feasible version of this proposition.

Obviously, if one is interested only in the variance (in period  $\tau$ ) of the transitory component of the  $i$ -th element of  $y$ , the  $i$ -th entry on the diagonal of  $J_{GG}\Omega J'_{GG}$  can be used.

Nevertheless, it is important to keep in mind that the derived confidence intervals are only valid for the chosen period  $\tau$  and not as confidence bands for the entire sample, since we cannot condition on the entire sample and still have random estimates. When we display our calculations in a form that resembles confidence bands for the time series, it is just done for convenience, since different readers may be interested in different periods.

### 3.3 Definition and representation of the Stock-Watson decomposition

In a standard formulation, and assuming a fixed initial value, the permanent SW components are given by

$$y_t^{permaSW} = y_0 + C\mu t + C \sum_{s=1}^t \varepsilon_s, \quad (10)$$

where  $C$  is the long-run moving-average impact matrix of reduced rank (which however is not directly of interest here). For the co-integrated VAR model the SW decomposition essentially yields the multivariate Beveridge-Nelson decomposition, i.e. the permanent component is a multivariate random walk. In contrast, the permanent component of the GG decomposition is autocorrelated in differences. This property of the SW decomposition implies an appealing interpretation: Given our knowledge at time  $t$ , only the SW transitory component of the time series is expected to change in the future (because it is expected to converge to its unconditional expectation, or in the demeaned case, to zero), so it is especially important for fore-

casting. Of course, the GG and SW permanent components only differ by stationary terms and are co-integrated, therefore they share the same long-run features.

Again following [Proietti \(1997\)](#) and [Hecq, Palm, and Urbain \(2000\)](#) the transitory SW component can be written as the sum of two terms,

$$y_t^{transSW} = \psi_t = \psi_{1t} + \psi_{2t}, \quad (11)$$

where the part  $\psi_{1t}$  represents the error-correcting movements of the system and is identical to the GG transitory component above, while the part  $\psi_{2t}$  are the remaining transitory movements of the system which do not contribute to the long-run equilibrium. For this latter part we need to define another lag polynomial if  $p > 1$ :  $B^*(L) = B_0^* + B_1^*L + \dots + B_{p-2}^*L^{p-2}$ , where  $B_j^* = \sum_{i=j+1}^{p-1} B_i$ . Then it can be represented as a distributed lag of the observable variables:

$$\psi_{2t} = -(I - P)Q^{-1}B^*(L)\Delta y_t \quad (12)$$

This second part remains to be demeaned as well, which can be achieved by using the known unconditional expectation of the differences:

$$E(\Delta y_t) = (I - P)Q^{-1}\mu \quad (13)$$

Using the abbreviation  $\mu^* \equiv (I - P)Q^{-1}\mu$  we can now write:

$$\begin{aligned}
\tilde{\Psi}_{2t} &= \Psi_{2t} - E(\Psi_{2t}) \\
&= -(I-P)Q^{-1}B^*(L)(\Delta y_t - \mu^*) \\
&= (-[I-P]Q^{-1})(B_0^*, -B_0^*\mu^*, B_1^*, -B_1^*\mu^*, \dots, B_{p-2}^*, -B_{p-2}^*\mu^*) \times \\
&\quad (\Delta y'_t, 1, \Delta y'_{t-1}, 1, \dots, \Delta y'_{t-p+2}, 1)' \\
&= (-[I-P]Q^{-1})(B_0^*, B_1^*, \dots, B_{p-2}^*, -B^*(1)\mu^*) \times \\
&\quad (\Delta y'_t, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+2}, 1)'
\end{aligned} \tag{14}$$

Then combining the two parts we have for the SW transitory component:

$$\begin{aligned}
\tilde{\Psi}_t &= \tilde{\Psi}_{1t} + \tilde{\Psi}_{2t} \\
&= (P, -[I-P]Q^{-1}[B_0^*, B_1^*, \dots, B_{p-2}^*], s_\mu) \times \\
&\quad (y'_t, \Delta y'_t, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+2}, 1)' \\
&\equiv S(y'_t, \Delta y'_t, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+2}, 1)',
\end{aligned} \tag{15}$$

where the last element relating to the constant term is given by

$$s_\mu = (Q^{-1}\alpha[\beta'Q^{-1}\alpha]^{-1}[\beta'Q^{-1}\alpha]^{-1}\beta' + [I-P]Q^{-1}B^*(1)[I-P])Q^{-1}\mu. \tag{16}$$

Note that also for the SW transitory component it is known how to test the hypothesis that it vanishes for a certain variable. In addition to the zero row of  $\alpha$  that was needed for the vanishing GG component, here the  $i$ -th rows of the various short-run coefficient matrices would also have to be zero. These restrictions essentially mean that the variable would be a strongly exogenous random walk. Again, for a given co-integration rank and super-consistently estimated co-integration co-

efficients, that would be a standard test problem.

### 3.4 The Delta method for the SW decomposition

The calculation of the uncertainty for the SW transitory component is analogous to the procedure for the GG component above. Again we can express the  $\tilde{\psi}_\tau$  (the demeaned overall transitory components) in period  $\tau \in \{p, \dots, T\}$  as a function of  $k$ , of the co-integration coefficients  $\beta$ , and of the data; the only difference now is that we have to condition on the lagged values as well,  $y_\tau, \dots, y_{\tau-p+1}$ :

$$\tilde{\psi}_\tau = f_{SW}(k; \beta, y_\tau, \dots, y_{\tau-p+1}) \quad (17)$$

The function  $f_{SW}$  is given by (15),  $f_{SW} = S(y'_t, \Delta y'_t, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+2}, 1)'$ . Let  $J_{SW} = \partial \tilde{\psi}_\tau / \partial k'$  be the Jacobian matrix of that function with respect to  $k$ , where the details are again provided in the appendix. We can state the estimation uncertainty of  $\tilde{\psi}_\tau$  similar to the one of the GG decomposition in proposition 1.

**Proposition 2.** *The conditional asymptotic distribution of the SW transitory component estimator is given by:*

$$\sqrt{T}(\hat{\psi}_\tau - \tilde{\psi}_\tau) \xrightarrow{d} N(0, J_{SW} \Omega J'_{SW}) \quad (18)$$

Again the result follows directly from applying the delta method, cf. the remarks on proposition 1, where now the conditioning data are given by  $y_\tau, \dots, y_{\tau-p+1}$ . The influence of these data on the actual estimates is either non-existent (if appropriate impulse dummies have been used in estimating the system), or it is asymptotically vanishing, see remark 2.

The result can be made feasible again by using the estimates  $\hat{k}$  and  $\hat{\Theta}$  for the construction of  $\hat{\Omega}$  and  $\hat{J}_{SW}$ . Of course, the interpretation remains only valid for a single chosen period.

## 4 The bootstrap method

The justification of the bootstrap in this case rests essentially on the same foundations as the delta method before. The underlying coefficients are freely varying (for a maintained co-integration rank  $r$ ), and the asymptotic distribution of the transitory components conditional on the data at a certain observation period  $\tau$  is well-behaved. Of course we hope that the bootstrap may yield some small-sample refinements over the asymptotic approximation by the delta method, for example by taking into account explicitly the variation of the co-integration coefficients estimates.

To be concrete, the distribution of the GG transitory component for the period of interest  $\tau$  can be simulated with the following algorithm. As a starting point we can use the standard estimates of (2).

1. Using the point estimates as the auxiliary data-generating process, simulate artificial data for the periods  $t = p \dots T$  by drawing from a suitable distribution describing the innovation process  $\varepsilon_t$ . This could either be a random draw from a fitted parametric distribution like a multivariate normal distribution with covariance matrix  $\widehat{\Theta}$  (and mean zero, of course), or re-sampling from the estimated residuals. We will use the observed values of  $y_t$  as the initial values of the artificial data in periods  $t = 0 \dots p - 1$ . Note that even though the resulting artificial data may be very different from the original data because it will have different underlying realizations of the stochastic trends, this does not affect the distribution of the transitory components.
2. Re-estimate the VECM using the same specification that was applied to the original data, but with the artificial data created in the previous step. Then record the estimates of  $\tilde{\psi}_{1\tau}$  as defined in equation (7), which means using the new estimated  $G$  coefficients of the current simulation run, but always employing the originally observed data  $(y'_\tau; 1)$ . Denote that estimate by  $\tilde{\psi}_{1\tau,w}$ ,

where  $w$  is a simulation index running from 1 to some sufficiently large integer  $W$ .

3. Repeat the previous two steps  $W$  times to get simulated distributions of (the estimate of)  $\tilde{\psi}_{1\tau}$ .
4. For the  $i$ -th variable calculate variants of the confidence intervals for the estimate of  $\tilde{\psi}_{1\tau}$  in the following two ways:

- (a) First we base the intervals directly on the distributions of  $\tilde{\psi}_{1\tau,w}$  over all  $w$  and construct a confidence interval using the empirical quantiles of the simulated distributions: with  $\gamma$  as the nominal coverage of the error band (1 minus the type-1 error) and the quantiles of  $\tilde{\psi}_{1\tau,w}$  given by  $\tilde{\psi}_{1\tau,(1-\gamma)/2}$  and  $\tilde{\psi}_{1\tau,(1+\gamma)/2}$ , the intervals are constructed as

$$[\tilde{\psi}_{1\tau,(1-\gamma)/2}, \tilde{\psi}_{1\tau,(1+\gamma)/2}]. \quad (19)$$

This construction is analogous to what [Sims and Zha \(1999\)](#) have called “other-percentile” bands in the slightly different context of impulse-response analysis, and they criticized their use as “clearly [amplifying] any bias present in the estimation procedure” (p. 1125).

- (b) Because of this criticism we also consider a Hall-type bootstrap, where the relevant distributions are given by  $\tilde{\psi}_{1\tau,w} - \tilde{\psi}_{1\tau}$ , i.e., for each variable and period the bootstrap realizations are corrected by the original point estimate.<sup>5</sup> Denoting the quantiles of these corrected distributions by  $(\tilde{\psi}_{1\tau,w} - \tilde{\psi}_{1\tau})_{(1-\gamma)/2}$  and  $(\tilde{\psi}_{1\tau,w} - \tilde{\psi}_{1\tau})_{(1+\gamma)/2}$ , the Hall-type error bands are given by

$$[\tilde{\psi}_{1\tau} - (\tilde{\psi}_{1\tau,w} - \tilde{\psi}_{1\tau})_{(1+\gamma)/2}, \tilde{\psi}_{1\tau} - (\tilde{\psi}_{1\tau,w} - \tilde{\psi}_{1\tau})_{(1-\gamma)/2}]. \quad (20)$$

---

<sup>5</sup>In order not to overload the notation, we do not formally distinguish here between the true transitory component (true of course conditional on period- $\tau$  data) and its original point estimate, because we hope it is clear from the context that only the estimate can be used here.



Note that the upper quantiles of the corrected distributions are used for the calculation of the lower error band margins, and vice versa. This “counter-acting swapping” serves to cancel out any bias of the estimation procedure.

Of course the bootstrap procedure can be simultaneously applied to all periods in the sample. However, we still do not get confidence “bands” because we cannot condition on the entire sample and do valid inference. As with the delta method, we can only derive valid confidence intervals for certain periods of interest.

For the SW transitory component the bootstrap method in this case is completely analogous to the GG case and to save space we will not repeat the details of the algorithm here. Essentially, the distribution of the  $G$  coefficients is replaced by that of the  $S$  coefficients, and of course the transitory component must be constructed using the extended data vector which includes lags, according to the formulas in section 3.3.

## 5 Illustration

For an illustration of how the methods work in practice we use a three-variable dataset inspired by the influential [King, Plosser, Stock, and Watson \(1991, KPSW\)](#) article dealing with stochastic trends in US business-cycle analysis. That is, we also use the quarterly variables (logs of) real consumption  $cons_t$ , real (gross) investment  $inv_t$ , and real output  $inc_t$ , but instead of their sample 1947-1988 we analyze more recent data spanning 1968q1-2010q2. We also let the series be tied together by two co-integrating relationships ( $r = 2$ ), such that any two of the three variables are co-integrated. KPSW propose to specify the co-integrating relationships according to economic theory as the “great ratios” of balanced growth, specifically  $cons - inc$  and  $inv - inc$ , but for the purposes of this illustration we will work with

freely estimated co-integration coefficients  $\beta$ .<sup>6</sup> No exogenous terms are included in the co-integration space, and the constant term is unrestricted to account for the deterministic long-run growth trend. As in KPSW we use eight lags ( $p = 8$ ) in this illustration.

For the transitory components we focus on excess output ( $y_t^{trans} = y_t - y_t^{perma}$ , the negative of the output gap); figure 1 shows the point estimates of the transitory components of both decompositions, GG and SW. Both estimates seem quite similar for this data –apart from fluctuations of the SW gap measure in the very short run– which may suggest the presence of common cyclical features (Proietti, 1997). The great recession of 2008-2010 is clearly visible as a large drop in excess output. In general we note that a falling excess output measure corresponds quite well to the NBER dating of the US recessions.

Before turning to the estimation uncertainty of the transitory components in specific periods it may be useful to briefly test whether the transitory output component is at all significant in this system. For the GG transitory component this check can be directly implemented as the standard test of the null hypothesis of a zero row in  $\alpha$  for the output equation. In our illustration here, this test yields the following result:  $P(\chi^2(2) > 11.19) = 0.0037$ , and thus the GG output gap is clearly significant in general, over the entire sample. Obviously, this finding automatically implies the significance of the SW output gap, since having a row of zeros in  $\alpha$  is also a necessary (but not sufficient) condition for a vanishing SW transitory component.

In the next step we calculate the confidence intervals for the output gap as given by the GG decomposition, shown in figure 2. All intervals have a nominal asymptotic coverage of 90%, and we have employed the described shortcut where the observation on which we condition is still included in the estimation sample, but its influence should be negligible compared to the rest of the sample. For the bootstraps

---

<sup>6</sup>These great ratios are actually not so great in terms of their stationarity properties in the subsample after the publication of KPSW. This is another reason to freely estimate the co-integration coefficients instead of imposing unit values.

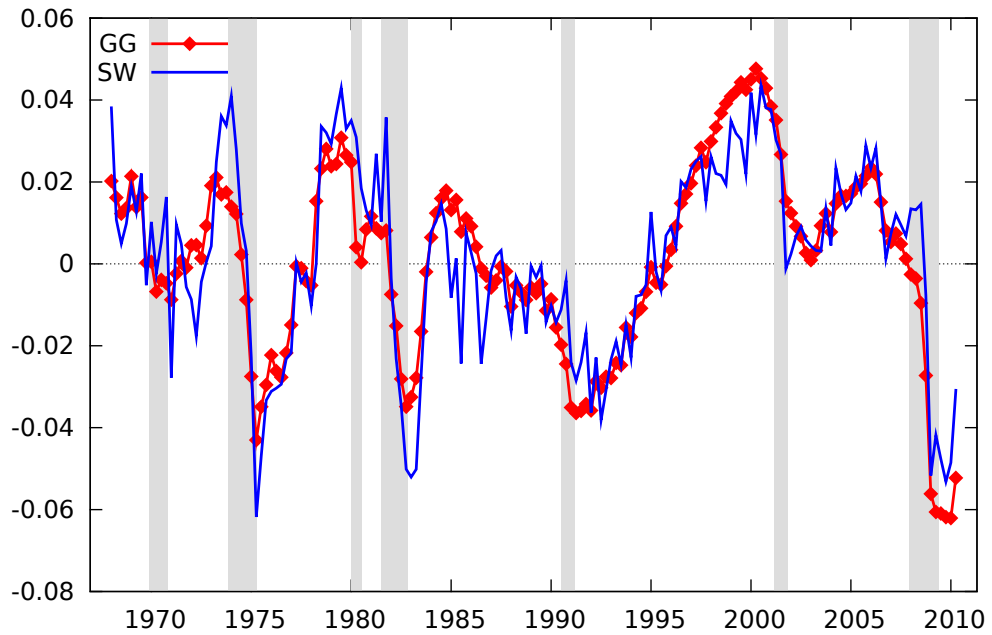


Figure 1: Estimated output gaps as transitory components of the GG and SW permanent-transitory decompositions. Shaded areas indicate NBER recession dating.

we re-sample from the estimated residuals.

First of all we notice that for most periods the intervals are quite similar. However, there are some exceptions; around 2005 for example the delta method intervals are tighter than their bootstrapped counterparts. And in the final observations for 2010 the Hall-type bootstrap intervals are shifted upwards somewhat in comparison to the naive bootstrap intervals (as well as compared to the delta method intervals, which are of course symmetric around the respective point estimates of the gap). Finally, whether the confidence intervals in general should be perceived as relatively wide or tight is probably a matter of taste. Nevertheless, while for the latest observation(s) the interval is quite wide indeed, the output gap interval estimate does not cover zero.

Finally, figure 3 displays the corresponding measures and calculations for the SW decomposition of the output series in the co-integrated system. Similar remarks as before apply concerning the comparison of the three different interval “series”.

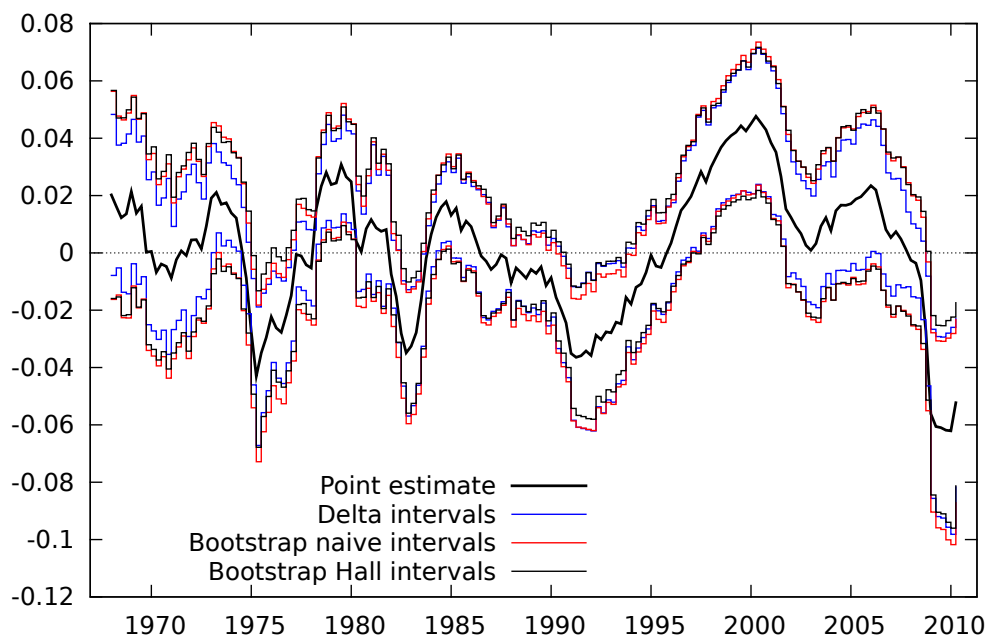


Figure 2: GG decomposition, confidence intervals for the output gap; displayed together for all periods in the sample for convenience, while the interpretation should be for a single period only.

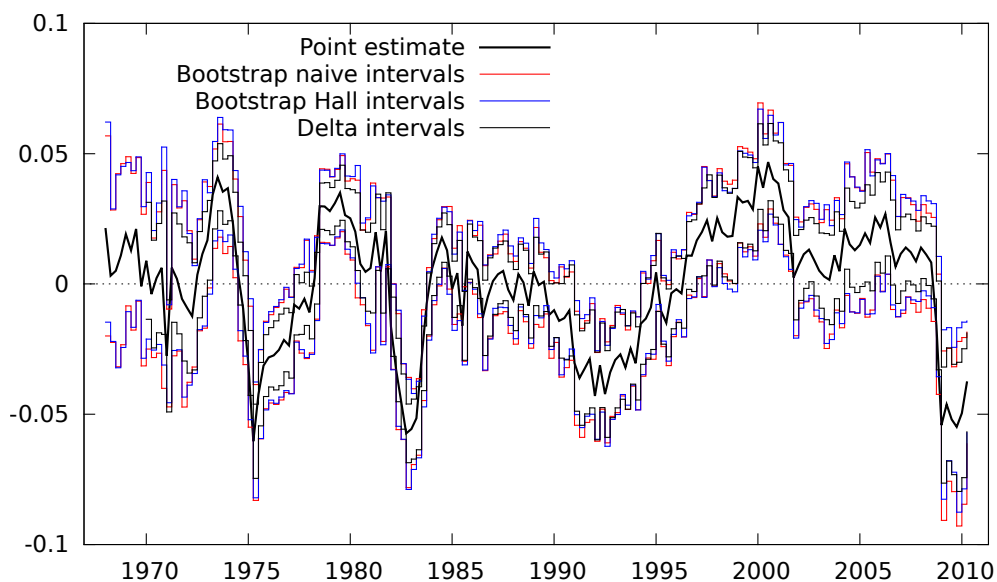


Figure 3: SW decomposition, confidence intervals for the output gap; displayed together for all periods in the sample for convenience, while the interpretation should be for a single period only.

## 6 Summary

While a permanent-transitory decomposition of non-stationary time series in a cointegrated system can always be mechanically calculated, it has not been clear until now if even the sign of the transitory component in the period of interest can be fully established with a sufficient degree of confidence, given the sampling uncertainty of the estimated coefficients. So far it has only been possible to test the overall significance of the transitory components for the entire sample.

Therefore, we have proposed an additional approach to assess the sampling uncertainty of widespread permanent-transitory decompositions, where we take as given the data constellations that are observed at the period of interest (possibly the latest observation period available). These measures provide additional information compared to the standard overall test results. For this conditional approach we have derived a delta-method and two bootstrap-based ways to quantify the estimation uncertainty of the Stock-Watson (common-trends-based) and Gonzalo-Granger (common-factor-based) decompositions.

In the empirical illustration we calculated the uncertainty of output gap estimates for the US. For example, it turned out that for the latest available observation (2010q2) the GG-based and SW-based output gap estimates have very wide confidence intervals (for 90% nominal coverage) which do not include zero, however.

## A Derivation of the Jacobians

In this appendix we provide a detailed derivation of the relevant Jacobian matrices needed for the delta method. The symbol  $\otimes$  denotes the Kronecker product.

First we address the Gonzalo-Granger transitory component  $\tilde{\psi}_t^{GG} = \tilde{\psi}_{1t}$ , followed by the second transitory component  $\tilde{\psi}_{2t}$ , which together form the Stock-Watson (-Proietti) transitory component  $\tilde{\psi}_t^{SW} = \tilde{\psi}_{1t} + \tilde{\psi}_{2t}$ . The short-run parameters of the model are collected in the following parameter matrix of dimension

$(n, (r + (p - 1)n + 1))$ :

$$Param = \begin{bmatrix} \alpha, & B_1, & \dots, & B_{p-1}, & \mu \end{bmatrix},$$

$(n,r) \quad (n,n) \qquad \qquad \qquad (n,1)$

and we define the parameter vector as

$$k = vec(Param) = [vec(\alpha)', vec(B_1)', \dots, vec(B_{p-1})', vec(\mu)']'.$$

## A.1 Useful rules

For convenience we reproduce some of the differentiation rules and other useful formulas from [Lütkepohl \(1997\)](#), the notation is unrelated to ours of the rest of the paper.

- 10.6.3 (5), p. 201, which is a quite general formula that can be used for several nested special cases. All matrices are in general functions of the vector  $x$ , the matrix  $U$  has  $q$  rows and the matrix  $W$  has  $r$  columns:

$$\begin{aligned} \frac{\partial vec(UY^{-1}VZ^{-1}W)}{\partial x'} &= (W'Z'^{-1}V'Y'^{-1} \otimes I_q) \frac{\partial vec(U)}{\partial x'} \\ &\quad - (W'Z'^{-1}V'Y'^{-1} \otimes UY^{-1}) \frac{\partial vec(Y)}{\partial x'} \\ &\quad + (W'Z'^{-1} \otimes UY^{-1}) \frac{\partial vec(V)}{\partial x'} \\ &\quad - (W'Z'^{-1} \otimes UY^{-1}VZ^{-1}) \frac{\partial vec(Z)}{\partial x'} \\ &\quad + (I_r \otimes UY^{-1}VZ^{-1}) \frac{\partial vec(W)}{\partial x'} \end{aligned}$$

- 10.4.4 (8), p. 188, where the matrix  $Y$  is a function of the vector  $x$ :

$$\frac{\partial vec(AYB + C)}{\partial x'} = (B' \otimes A) \frac{\partial vec(Y)}{\partial x'}$$

- 7.2 (5), p. 97, where the matrix  $A$  has  $m$  rows and  $n$  columns, and  $B$  is  $(n, p)$ :

$$\text{vec}(AB) = (B' \otimes I_m)\text{vec}(A) = (B' \otimes A)\text{vec}(I_n) = (I_p \otimes A)\text{vec}(B)$$

- 2.4(13) for the special case of two column vectors  $x$  and  $y$  (which also implies  $\text{vec}(x \otimes y) = x \otimes y$ ):

$$\text{vec}(yx') = x \otimes y$$

- 10.4.1 (3), p.183:

$$\frac{\partial \text{vec}(AXB)}{\partial \text{vec}(X)'} = B' \otimes A$$

- 10.6.3 (1), p.200:

$$\frac{\partial \text{vec}(Y^{-1})}{\partial x'} = -(Y'^{-1} \otimes Y^{-1}) \frac{\partial \text{vec}(Y)}{\partial x'}$$

## A.2 The Gonzalo-Granger component

The component  $\tilde{\psi}_{1\tau} = G(k; \beta)(y'_{\tau}, 1)'$  is a function of  $k$  via  $G$ , conditional on the data  $y_{\tau}$ , where  $G(k; \beta) = Q^{-1}\alpha[\beta'Q^{-1}\alpha]^{-1}(\beta', [\beta'Q^{-1}\alpha]^{-1}\beta'Q^{-1}\mu)$  is a partitioned matrix, corresponding to the partitioning of  $(y'_t, 1)$ . (The first part of  $G$ ,  $Q^{-1}\alpha[\beta'Q^{-1}\alpha]^{-1}\beta'$ , is Proietti's  $P$  matrix.)

The covariance matrix of  $k$  is directly given from the system estimation, and to infer the covariance matrix of  $\tilde{\psi}_{1t}$  by the delta method we need the Jacobian of this vector-to-vector mapping, i.e.:

$$\frac{\partial \text{vec}(\tilde{\psi}_{1t})}{\partial \text{vec}(k)'} = \frac{\partial \tilde{\psi}_{1t}}{\partial k'}$$

We first formulate some of the parameters as explicit matrix products. For  $\alpha$  we have

$$\alpha = Param \times \begin{bmatrix} I_r \\ \mathbf{0}_{((p-1)n,r)} \\ \mathbf{0}_{(1,r)} \end{bmatrix},$$

and thus:

$$\begin{aligned} vec(\alpha) &\stackrel{7.2(5)}{=} ([I_r, \mathbf{0}_{(r,(p-1)n+1)}] \otimes I_n) k \\ &= [I_{nr}, \mathbf{0}_{(nr,n^2(p-1)+n)}] k \end{aligned} \quad (21)$$

Next,  $B_i$  is given as:

$$B_i = Param \times \begin{bmatrix} \mathbf{0}_{(r,n)} \\ \mathbf{0}_{((i-1)n,n)} \\ I_n \\ \mathbf{0}_{((p-1-i)n,n)} \\ \mathbf{0}_{(1,n)} \end{bmatrix},$$

which implies:

$$\begin{aligned} vec(B_i) &\stackrel{7.2(5)}{=} ([\mathbf{0}_{(n,r)}, \dots, I_n, \dots, \mathbf{0}_{(n,1)}] \otimes I_n) k \\ &= [\mathbf{0}_{(n^2,nr)}, \mathbf{0}_{(n^2,n^2(i-1))}, I_{n^2}, \mathbf{0}_{(n^2,n^2(p-1-i))}, \mathbf{0}_{(n^2,n)}] k \end{aligned} \quad (22)$$

In the same way, the estimated constant term  $\mu$  can be expressed as:

$$\mu = Param \times \begin{bmatrix} \mathbf{0}_{(r+(p-1)n,1)} \\ 1 \end{bmatrix},$$



and also  $\text{vec}(\mu) = \mu$ , so:

$$\begin{aligned} \mu &\stackrel{7.2(5)}{=} ([\mathbf{0}_{(1,r+(p-1)n)}, \mathbf{1}] \otimes I_n) k \\ &= [\mathbf{0}_{(n,nr+n^2(p-1))}, I_n] k \end{aligned} \quad (23)$$

Then, using the abbreviation  $\tilde{\mu} \equiv [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1} \mu$  we can also re-express the partitioned matrix  $(\beta', [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1} \mu) \equiv (\beta', \tilde{\mu})$  as a product/sum. Since the second part is a column vector  $(r, 1)$ ,  $\text{vec}(\tilde{\mu}) = \tilde{\mu}$ , first it is clear that simply

$$\text{vec}(\beta'; \tilde{\mu}) = \begin{bmatrix} \text{vec}(\beta') \\ \tilde{\mu} \end{bmatrix} = \begin{bmatrix} \text{vec}(\beta') \\ \mathbf{0}_{(r,1)} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{(rn,1)} \\ \tilde{\mu} \end{bmatrix}.$$

However, it is useful to express especially the second part which depends on  $k$  as a matrix product to apply the formulas for derivatives. We have  $(\mathbf{0}_{(r,n)}, \tilde{\mu}) = (\mathbf{0}_{(1,n)}, \mathbf{1}) \otimes \tilde{\mu}$ , and after stacking we get:

$$\begin{bmatrix} \mathbf{0}_{(rn,1)} \\ \tilde{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(n,1)} \\ 1 \end{bmatrix} \otimes \tilde{\mu} \stackrel{2.4(13)}{=} \text{vec}(\tilde{\mu} [\mathbf{0}_{(1,n)}, \mathbf{1}])$$

So altogether we can use

$$(\beta', \tilde{\mu}) = (\beta', \mathbf{0}_{(r,1)}) + (\mathbf{0}_{(1,n)}, \mathbf{1}) \otimes \tilde{\mu}, \quad (24)$$

or in stacked form:

$$\text{vec}(\beta', \tilde{\mu}) = \begin{bmatrix} \text{vec}(\beta') \\ \mathbf{0}_{(r,1)} \end{bmatrix} + \text{vec}(\tilde{\mu} [\mathbf{0}_{(1,n)}, \mathbf{1}]) \quad (25)$$

Now we can calculate some derivatives of smaller parts:

$$\begin{aligned}
\frac{\partial \text{vec}(\alpha)}{\partial k'} &\stackrel{(21)}{=} \frac{\partial \left[ I_{nr}, \mathbf{0}_{(nr, n^2(p-1)+n)} \right] k}{\partial k'} \\
&= \left[ I_{nr}, \mathbf{0}_{(nr, n^2(p-1)+n)} \right]
\end{aligned} \tag{26}$$

With this result, and remembering that we treat  $\beta$  as asymptotically fixed, the derivative of the product  $\alpha\beta'$  is given by:

$$\begin{aligned}
\frac{\partial \text{vec}(\alpha\beta')}{\partial k'} &\stackrel{7.2(5)}{=} (\beta \otimes I_n) \frac{\partial \text{vec}(\alpha)}{\partial k'} \\
&\stackrel{(26)}{=} (\beta \otimes I_n) \left[ I_{nr}, \mathbf{0}_{(nr, n^2(p-1)+n)} \right]
\end{aligned} \tag{27}$$

Turning to the derivative of the sum of the  $B_i$  coefficients, we obtain:

$$\begin{aligned}
\sum_{i=1}^{p-1} \frac{\partial \text{vec}(B_i)}{\partial k'} &\stackrel{(22)}{=} \sum \frac{\partial \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{0}_{(n^2, n^2(i-1))}, I_{n^2}, \mathbf{0}_{(n^2, n^2(p-1-i))}, \mathbf{0}_{(n^2, n)} \right] k}{\partial k'} \\
&\stackrel{(10.4.1(3))}{=} \sum \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{0}_{(n^2, n^2(i-1))}, I_{n^2}, \mathbf{0}_{(n^2, n^2(p-1-i))}, \mathbf{0}_{(n^2, n)} \right] \\
&= \left[ \mathbf{0}_{(n^2, nr)}, I_{n^2}, \dots, I_{n^2}, \mathbf{0}_{(n^2, n)} \right] \\
&= \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{1}_{(1, p-1)} \otimes I_{n^2}, \mathbf{0}_{(n^2, n)} \right]
\end{aligned} \tag{28}$$

Putting this (28 and 27) together yields the derivative of  $Q = B(1) - \alpha\beta'$ :

$$\begin{aligned}
\frac{\partial \text{vec}(Q)}{\partial k'} &= \frac{\partial \text{vec}(B(1) - \alpha\beta')}{\partial k'} \\
&= \frac{\partial \text{vec}(I - \sum B_i - \alpha\beta')}{\partial k'} \\
&= - \sum \frac{\partial \text{vec}(B_i)}{\partial k'} - \frac{\partial \text{vec}(\alpha\beta')}{\partial k'} \\
&= - \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{1}_{(1, p-1)} \otimes I_{n^2}, \mathbf{0}_{(n^2, n)} \right] \\
&\quad - (\beta \otimes I_n) \left[ I_{nr}, \mathbf{0}_{(nr, n^2(p-1)+n)} \right]
\end{aligned} \tag{29}$$

And we can build upon this to also obtain the derivative of  $Q^{-1}$ :

$$\frac{\partial \text{vec}(Q^{-1})}{\partial k'} \stackrel{10.6.3(1)}{=} - (Q'^{-1} \otimes Q^{-1}) \frac{\partial \text{vec}(Q)}{\partial k'} \quad (30)$$

Now the derivative of (the vector)  $\mu$  is quite straightforward from (23):

$$\frac{\partial \mu}{\partial k'} = \left[ 0_{(n, nr + n^2(p-1))}, I_n \right] \quad (31)$$

Next we provide the derivative of  $\beta' Q^{-1} \alpha$ , where we use formula 10.6.3(5) (where we treat  $\beta' \triangleq U$ ,  $Q \triangleq Y$ ,  $\alpha \triangleq V$ , and the additional matrices  $Z$  and  $W$  are identity matrices in this case):

$$\begin{aligned} \frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'} &= 0 - (\alpha' Q'^{-1} \otimes \beta' Q^{-1}) \frac{\partial \text{vec}(Q)}{\partial k'} + (I_r \otimes \beta' Q^{-1}) \frac{\partial \text{vec}(\alpha)}{\partial k'} - 0 + 0 \\ &= (\alpha' Q'^{-1} \otimes \beta' Q^{-1}) \times \\ &\quad \left( \left[ 0_{(n^2, nr)}, 1_{(1, p-1)} \otimes I_{n^2}, 0_{(n^2, n)} \right] + (\beta \otimes I_n) \left[ I_{nr}, 0_{(nr, n^2(p-1)+n)} \right] \right) \\ &\quad + (I_r \otimes \beta' Q^{-1}) \left[ I_{nr}, 0_{(nr, n^2(p-1)+n)} \right] \\ &= (\alpha' Q'^{-1} \otimes \beta' Q^{-1}) \left[ 0_{(n^2, nr)}, 1_{(1, p-1)} \otimes I_{n^2}, 0_{(n^2, n)} \right] + \\ &\quad + ((\alpha' Q'^{-1} \beta \otimes \beta' Q^{-1}) + (I_r \otimes \beta' Q^{-1})) \left[ I_{nr}, 0_{(nr, n^2(p-1)+n)} \right] \\ &= (\alpha' Q'^{-1} \otimes \beta' Q^{-1}) \left[ 0_{(n^2, nr)}, 1_{(1, p-1)} \otimes I_{n^2}, 0_{(n^2, n)} \right] + \\ &\quad + ((\alpha' Q'^{-1} \beta + I_r) \otimes \beta' Q^{-1}) \left[ I_{nr}, 0_{(nr, n^2(p-1)+n)} \right] \end{aligned} \quad (32)$$

Using these previous intermediate results, and repeating the expression for the relevant transitory component,

$$\tilde{\Psi}_{1\tau} = Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} (\beta', \tilde{\mu}) \begin{bmatrix} y_\tau \\ 1 \end{bmatrix},$$

we apply formula 10.6.3(5) by setting  $U \triangleq Q^{-1}$ ,  $V \triangleq \alpha$ ,  $Z^{-1} \triangleq [\beta' Q^{-1} \alpha]^{-1}$ , and  $W$  as the rest of the expression (the matrix  $Y$  of the formula is an identity matrix here).

Obviously, this is where the conditioning data  $y_\tau$  enter. Thus:

$$\begin{aligned}
\frac{\partial \tilde{\psi}_{1\tau}}{\partial k'} &= \left\{ \left( (y'_\tau, 1) (\beta', \tilde{\mu})' [\beta' Q^{-1} \alpha]'^{-1} \alpha' \right) \otimes I_n \right\} \frac{\partial \text{vec}(Q^{-1})}{\partial k'} \\
&\quad - 0 \\
&\quad + \left\{ (y'_\tau, 1) (\beta', \tilde{\mu})' [\beta' Q^{-1} \alpha]'^{-1} \otimes Q^{-1} \right\} \frac{\partial \text{vec}(\alpha)}{\partial k'} \\
&\quad - \left\{ (y'_\tau, 1) (\beta', \tilde{\mu})' [\beta' Q^{-1} \alpha]'^{-1} \otimes Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} \right\} \frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'} \\
&\quad + Q^{-1} \alpha [\beta' Q^{-1} \alpha]^{-1} \frac{\partial \text{vec} \left( (\beta', \tilde{\mu}) [y'_\tau, 1]' \right)}{\partial k'} \tag{33}
\end{aligned}$$

where the embedded derivatives  $\frac{\partial \text{vec}(Q^{-1})}{\partial k'}$ ,  $\frac{\partial \text{vec}(\alpha)}{\partial k'}$ , and  $\frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'}$  have been solved before in (30), (26), and (32), and with respect to the last line we also notice that  $(\beta', \tilde{\mu}) (y'_\tau, 1)' = \beta' y_\tau + \tilde{\mu}$ , and since  $\beta' y_\tau$  does not depend on  $k$ , we only need to determine the derivative  $\partial \tilde{\mu} / \partial k'$ . Here we again apply 10.6.3(5) by setting  $U = I, Y^{-1} = [\beta' Q^{-1} \alpha]^{-1}, V = \beta', Z^{-1} = Q^{-1}, W = \mu$ , which yields:

$$\begin{aligned}
\frac{\partial \text{vec}(\tilde{\mu})}{\partial k'} &= \frac{\partial \tilde{\mu}}{\partial k'} = 0 \\
&\quad - (\mu' Q^{-1} \beta (\beta' Q^{-1} \alpha)'^{-1} \otimes [\beta' Q^{-1} \alpha]^{-1}) \frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'} \\
&\quad + 0 \\
&\quad - (\mu' Q^{-1} \otimes [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1}) \frac{\partial \text{vec}(Q)}{\partial k'} \\
&\quad + [\beta' Q^{-1} \alpha]^{-1} \beta' Q^{-1} \frac{\partial \mu}{\partial k'} \tag{34}
\end{aligned}$$

The embedded derivatives  $\frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'}$ ,  $\frac{\partial \text{vec}(Q)}{\partial k'}$ ,  $\frac{\partial \mu}{\partial k'}$  are given in equations (32), (29), and (31). Inserting these expressions we obtain an operational version of equation (34) for use in (33).

Most of the computations only have to be performed once (per estimation sample), and for each analyzed observation  $\tau$  only the matrix multiplications and addi-

tions in (33) have to be calculated that involve the conditioning data  $y_\tau$ .

### A.3 The Stock-Watson component

As  $\tilde{\psi}_\tau = \tilde{\psi}_{1,\tau} + \tilde{\psi}_{2,\tau}$ , where the first part is the Gonzalo-Granger component for which we have already found the derivative, only  $\partial \tilde{\psi}_{2,\tau} / \partial k'$  remains to be computed. Since  $\tilde{\psi}_{2,\tau} = \psi_{2,\tau} - E(\psi_{2,\tau})$ , we can separately analyze  $\partial \psi_{2,\tau} / \partial k'$  and  $\partial E(\psi_{2,\tau}) / \partial k'$ , and subtract them afterwards.

As  $B_j^* = \sum_{i=j+1}^{p-1} B_i$  is involved ( $j = 0..p-2$ ), we can use  $\partial \text{vec}(B_i) / \partial k'$  from (28), where it is implicitly used, to obtain:

$$\begin{aligned} \frac{\partial \text{vec}(B_j^*)}{\partial k'} &= \frac{\partial \text{vec}\left(\sum_{i=j+1}^{p-1} B_i\right)}{\partial k'} = \sum_{i=j+1}^{p-1} \frac{\partial \text{vec}(B_i)}{\partial k'} \\ &= \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{0}_{(1, j)} \otimes I_{n^2}, \mathbf{1}_{(1, p-1-j)} \otimes I_{n^2}, \mathbf{0}_{(n^2, n)} \right], \end{aligned} \quad (35)$$

where the  $\mathbf{0}_{(1, j)} \otimes I_{n^2}$  element comes from the lacking  $B_1 \dots B_j$  components.

#### A.3.1 The part $\partial \psi_{2,\tau} / \partial k'$

The non-demeaned additional component relating to the short-run dynamics in the differences is given by

$$\psi_{2,\tau} = -(I - P)Q^{-1}B^*(L)\Delta y_\tau,$$

where  $B^*(L) = B_0^* + B_1^*L + \dots + B_{p-2}^*L^{p-2}$ , and as before  $P = Q^{-1}\alpha(\beta'Q^{-1}\alpha)^{-1}\beta'$ .

This can also be written with explicit data vectors:

$$\psi_{2,\tau} = -(I - P)Q^{-1} \begin{bmatrix} B_0^* & \dots & B_{p-2}^* \end{bmatrix} \begin{bmatrix} \Delta y_\tau \\ \Delta y_{\tau-1} \\ \vdots \\ \Delta y_{\tau-(p-2)} \end{bmatrix}$$

In order to differentiate this component, we apply formula 10.6.3(5) again by setting:  $U \hat{=} -(I_n - P) = P - I_n$ ,  $Y^{-1} \hat{=} Q^{-1}$ ,  $V \hat{=} [B_0^*, \dots, B_{p-2}^*]$ ,  $Z^{-1} \hat{=} I_{n(p-1)}$ , and  $W \hat{=} [\Delta y'_{\tau}, \dots, \Delta y'_{\tau-(p-2)}]'$ . We get:

$$\begin{aligned}
\frac{\partial \psi_{2,\tau}}{\partial k'} &= \left( [\Delta y'_{\tau}, \dots, \Delta y'_{\tau-(p-2)}] [B_0^*, \dots, B_{p-2}^*]' Q'^{-1} \otimes I_n \right) \frac{\partial \text{vec}(P - I_n)}{\partial k'} \\
&\quad - \left( [\Delta y'_{\tau}, \dots, \Delta y'_{\tau-(p-2)}] [B_0^*, \dots, B_{p-2}^*]' Q'^{-1} \otimes (P - I)Q^{-1} \right) \frac{\partial \text{vec}(Q)}{\partial k'} \\
&\quad + \left( [\Delta y'_{\tau}, \dots, \Delta y'_{\tau-(p-2)}] \otimes (P - I)Q^{-1} \right) \frac{\partial \text{vec} [B_0^*, \dots, B_{p-2}^*]}{\partial k'} \quad (36) \\
&\quad - 0 \\
&\quad + 0
\end{aligned}$$

The derivative in the second line (dealing with  $Q$ ) is known from equ. (29). The derivative in the first line is equal to  $\partial \text{vec}(P)/\partial k' = \partial \text{vec}(Q^{-1} \alpha (\beta' Q^{-1} \alpha)^{-1} \beta') / \partial k'$ , which is of course very similar to the result in equ. (33), except that  $(\beta', \tilde{\mu})(y'_{\tau}, 1)'$  is replaced by just  $\beta'$  at the end. Therefore:

$$\begin{aligned}
\frac{\partial \text{vec}(P)}{\partial k'} &= \left( \beta (\beta' Q^{-1} \alpha)^{-1} \alpha' \otimes I_n \right) \frac{\partial \text{vec}(Q^{-1})}{\partial k'} \\
&\quad + \left( \beta (\beta' Q^{-1} \alpha)^{-1} \otimes Q^{-1} \right) \frac{\partial \text{vec}(\alpha)}{\partial k'} \\
&\quad - \left( \beta (\beta' Q^{-1} \alpha)^{-1} \otimes Q^{-1} \alpha (\beta' Q^{-1} \alpha)^{-1} \right) \frac{\partial \text{vec}(\beta' Q^{-1} \alpha)}{\partial k'} \quad (37)
\end{aligned}$$

All the embedded derivatives of this expression are already known, see (30), (26), and (32).

Finally we address the third line of (36). First we note that we can re-express

$[B_0^*, \dots, B_{p-2}^*]$  in stacked form:

$$\begin{aligned}
\text{vec} [B_0^*, \dots, B_{p-2}^*] &= \begin{bmatrix} 1 \\ \mathbf{0}_{(p-2,1)} \end{bmatrix} \otimes \text{vec}(B_0^*) + \dots + \begin{bmatrix} \mathbf{0}_{(p-2,1)} \\ 1 \end{bmatrix} \otimes \text{vec}(B_{p-2}^*) \\
&= \sum_{j=0}^{p-2} \left( \begin{bmatrix} \mathbf{0}_{(j,1)} \\ 1 \\ \mathbf{0}_{(p-2-j,1)} \end{bmatrix} \otimes \text{vec}(B_j^*) \right) \\
&= \sum_{j=0}^{p-2} \text{vec} (\text{vec}(B_j^*) [\mathbf{0}_{(1,j)}, 1, \mathbf{0}_{(1,p-2-j)}])
\end{aligned}$$

where in the last line we have exploited the formula 2.4(13) for column vectors.

Now we can find the derivative of  $\text{vec} [B_0^*, \dots, B_{p-2}^*]$  by applying formula 10.4.4(8):

$$\frac{\partial \text{vec} [B_0^*, \dots, B_{p-2}^*]}{\partial k'} = \sum_{j=0}^{p-2} \left( \begin{bmatrix} \mathbf{0}_{(j,1)} \\ 1 \\ \mathbf{0}_{(p-2-j,1)} \end{bmatrix} \otimes I_{n^2} \right) \frac{\partial \text{vec}(\text{vec}(B_j^*))}{\partial k'}$$

Of course,  $\frac{\partial \text{vec}(\text{vec}(B_j^*))}{\partial k'} = \frac{\partial \text{vec}(B_j^*)}{\partial k'}$ , which we know already from (35), and after

inserting that we get:

$$\begin{aligned}
\frac{\partial \text{vec} [B_0^*, \dots, B_{p-2}^*]}{\partial k'} &= \sum_{j=0}^{p-2} \left( \begin{bmatrix} \mathbf{0}_{(j,1)} \\ 1 \\ \mathbf{0}_{(p-2-j,1)} \end{bmatrix} \otimes I_{n^2} \right) \times \\
&\quad \left[ \mathbf{0}_{(n^2, nr)}, \mathbf{0}_{(1,j)} \otimes I_{n^2}, \mathbf{1}_{(1,p-1-j)} \otimes I_{n^2}, \mathbf{0}_{(n^2, n)} \right]
\end{aligned} \tag{38}$$

Thus we have provided all needed embedded derivatives to make equ. (36) and thus  $\partial \psi_{2,\tau} / \partial k'$  operational. Again, most of these matrix computations only have to be performed once per sample and can be reused with new constellations of conditioning data.

### A.3.2 The part $\partial E(\psi_{2,\tau})/\partial k'$

We still need to consider the de-meaning term:

$$E(\psi_{2,\tau}) = -(I-P)Q^{-1}B^*(1)(I-P)Q^{-1}\mu$$

and sometimes we also use the abbreviation  $\mu^* = (I-P)Q^{-1}\mu$ .

We again apply formula 10.6.3(5) with these correspondences:  $U \triangleq -(I-P)$ ,  $Y^{-1} \triangleq Q^{-1}$ ,  $V \triangleq B^*(1)(I-P)$ ,  $Z^{-1} \triangleq Q^{-1}$ ,  $W \triangleq \mu$ . This yields:

$$\begin{aligned} \frac{\partial \text{vec}(E(\psi_{2,\tau}))}{\partial k'} &= \left( \mu' Q'^{-1} (I_n - P)' B^*(1)' Q'^{-1} \otimes I_n \right) \frac{\partial \text{vec}(P - I_n)}{\partial k'} \\ &\quad - \left( \mu' Q'^{-1} (I_n - P)' B^*(1)' Q'^{-1} \otimes (P - I_n) Q^{-1} \right) \frac{\partial \text{vec}(Q)}{\partial k'} \\ &\quad + \left( \mu' Q'^{-1} \otimes (P - I_n) Q^{-1} \right) \frac{\partial \text{vec}(B^*(1)(I_n - P))}{\partial k'} \\ &\quad - \left( \mu' Q'^{-1} \otimes (P - I_n) Q^{-1} B^*(1)(I_n - P) Q^{-1} \right) \frac{\partial \text{vec}(Q)}{\partial k'} \\ &\quad + \left( (P - I_n) Q^{-1} B^*(1)(I_n - P) Q^{-1} \right) \frac{\partial \mu}{\partial k'} \end{aligned} \quad (39)$$

Most of the embedded derivatives here are already known, see (37), (29), and (31), but the third line needs to be addressed. By applying yet again formula 10.6.3(5) in a very special case ( $U \triangleq B^*(1)$  and  $V \triangleq I_n - P$ , the rest are identity matrices), we get:

$$\begin{aligned} \frac{\partial \text{vec}(B^*(1)(I - P))}{\partial k'} &= \left( (I_n - P)' \otimes I_n \right) \frac{\partial \text{vec}(B^*(1))}{\partial k'} \\ &\quad + \left( I_n \otimes B^*(1) \right) \frac{\partial \text{vec}(I_n - P)}{\partial k'} \end{aligned} \quad (40)$$

Here in turn we just need to determine the derivative in the first line; for the embedded derivative in the second line we have  $\partial \text{vec}(I - P)/\partial k' = -\partial \text{vec}(P)/\partial k'$ , which is known (up to the sign) from equ. (37). Because  $B^*(1) = \sum_{j=0}^{p-2} B_j^*$ , we



simply have:

$$\begin{aligned}
\frac{\partial \text{vec}(B^*(1))}{\partial k'} &= \sum_{j=0}^{p-2} \frac{\partial \text{vec}(B_j^*)}{\partial k'} \\
&= \left[ 0_{(n^2, nr)}, I_{n^2}, 2I_{n^2}, \dots, (p-1)I_{n^2}, 0_{(n^2, n)} \right] \\
&= \left[ 0_{(n^2, nr)}, [1, 2, \dots, p-1] \otimes I_{n^2}, 0_{(n^2, n)} \right] \quad (41)
\end{aligned}$$

where equ. (35) has been used.

This completes the derivation of the analytical Jacobians for the delta method as used in the paper.

## References

- COTTRELL, A., AND R. LUCCHETTI (2010): *Gretl User's Guide* Version 1.9; available at <http://gretl.sourceforge.net/#man>.
- GONZALO, J., AND C. GRANGER (1995): "Estimation of Common Long-Memory Components in Cointegrated Systems," *Journal of Business and Economic Statistics*, 13(1), 27–35.
- HARVEY, A. C., AND T. PROIETTI (eds.) (2005): *Readings in Unobserved Components Models*. Oxford University Press.
- HARVEY, A. C., AND N. G. SHEPHARD (1993): "Structural Time Series Models," in *Handbook of Statistics*, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod, vol. 11 (Econometrics). North-Holland.
- HECQ, A., F. C. PALM, AND J.-P. URBAIN (2000): "Permanent-transitory Decomposition in VAR Models with Cointegration and Common Cycles," *Oxford Bulletin of Economics and Statistics*, 62(4), 511–532.
- KING, R. G., C. I. PLOSSER, J. H. STOCK, AND M. W. WATSON (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81(4), 819–840.
- LÜTKEPOHL, H. (1997): *Handbook of Matrices*. Wiley.

- PARUOLO, P. (1997): "Asymptotic inference on the moving average impact matrix in cointegrated  $I(1)$  VAR systems," *Econometric Theory*, 13, 79–118.
- PROIETTI, T. (1997): "Short-Run Dynamics in Cointegrated Systems," *Oxford Bulletin of Economics and Statistics*, 59(3), 405–422.
- SIMS, C. A., AND T. ZHA (1999): "Error Bands For Impulse Responses," *Econometrica*, 67(5), 1113–1155.
- STOCK, J. H., AND M. W. WATSON (1988): "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097–1107.