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

by Sven Schreiber†

this version: August 2016

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 Jï¿œrg Breitung, Neil Ericsson, Helmut Lï¿œtkepohl, and Sï¿œren 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 Panthï¿œon-Sorbonne CNRS, where parts of the early research for this paper were carried out. He dedicates this work to the late Jï¿œrgen Wolters who introduced him to co-integration analysis. The numeric calculations have been produced with gretl (Cottrell and Lucchetti, 2013), the code is available as a gretl function package.

†Macroeconomic Policy Institute (IMK) at Hans Boeckler Foundation, and Free University Berlin. E-mail: mail@sven.schreiber.name, fax: +49(0)211-7778-4596, address: Hans-Bï¿œckler-Stiftung/IMK, Hans-Bï¿œckler-Str. 39, D-40476 Dï¿œsseldorf, Germany.

1 Introduction

In this paper we propose 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 finite-order 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 fairly flexible tool. Chang, Miller, and Park (2009) and Chang, Jiang, and Park (2012) develop rigorous inference for Kalman filter based estimation of transitory and permanent components of co-integrated time series, but proposition 4.1 of the latter paper clarifies that their chosen model class does not contain finite-order VARs or VECMs as analyzed here. Their setup implies a VARMA reduced-form representation. In contrast, for VECM based trend and cycle measures so far there is no way to quantify the estimation uncertainty for a given period of interest. It has only been possible to test the existence of a transitory component based on the whole sample (on average).[1] But even if a variable contains a transitory component, that component will be near zero quite

[1]As stressed by Oh, Zivot, and Creal (2008), the way how Morley (2002) casts the VECM in another state-space form to extract a permanent component does not deliver uncertainty measures.

often because it is stationary and thus mean-reverting. The aim of this paper is to provide tools to assess the estimation uncertainty around the GG and SW measures at a given time in history, for example whether the estimated transitory component is statistically indistinguishable from zero.

The general idea is to condition on the data constellation prevailing in a single period of interest (say, τ) and to focus on the parameter estimation uncertainty arising from the estimation sample. The result will be confidence intervals around the estimated permanent and transitory components of the variables in a single observation period of interest. For example, the final observation at the current edge of the sample would be interesting in a macroeconomic application where the transitory component of the output variable is interpreted as the estimated output gap. If the confidence interval around the estimated output gap did not include zero a natural interpretation would be a statistically significant finding of an overheating or recessionary economy, with potential implications for policy makers.

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 will be addressed. 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, and in Section 5 we investigate the performance of the various methods in a simulation study. In Section 6 both approaches are applied to a three-variable dataset inspired by the influential work of King, Plosser, Stock, and Watson (1991), but covering more recent data. Section 7 summarizes.

2 Framework and assumptions

Consider a standard n-dimensional VAR with p lags:

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + \mu + \varepsilon_t, \qquad t = p, \ldots, T \tag{1}$$

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

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

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 α and β). 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 μ 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 ε_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,[2] they facilitate our analytical treatment in this paper.

Assumption 3. *The co-integration rank r, $n > r > 0$, and the lag order p are taken to be given and fixed.*

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.

[2] 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.)

Assumption 4. *The co-integration coefficients β 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 β, not on α, 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 $(\alpha, B_1, ..., B_{p-1}, \mu)$ and stack the coefficients in the vector $k = vec(\alpha, B_1, ..., B_{p-1}, \mu)$; this vector has $nr + n^2(p-1) + n$ elements that are freely varying.[3] Note that β is not included here because its estimate will be treated as asymptotically fixed given its higher convergence rate (T instead of $\sqrt{T}$), 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.[4]

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 Ω 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

[3] 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.

[4] This also applies to the $\alpha_\perp$-directions of the constant term, see Paruolo (1997).

estimates $\hat{k}$ is identical whether or not some observations are removed from the sample. This means that for practical purposes we can use the whole available sample for parameter estimation, even though we will condition on the data in the period of interest τ. The slightly more complicated alternative would be to exclude the data related to period τ from the likelihood function, e.g. through the inclusion of appropriate impulse dummies in the estimation.

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. 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 optimization methods such as those used for the estimation of state-space models.

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

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

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

$$P = Q^{-1} \alpha \left[\beta' Q^{-1} \alpha \right]^{-1} \beta' \tag{4}$$

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:

$$y_t^{transGG} = \psi_{1t} = P y_t \tag{5}$$

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:

$$\mathsf{E}(\beta' y_t) = -\left(\beta' Q^{-1}\alpha\right)^{-1}\beta' Q^{-1}\mu, \tag{6}$$

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

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

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 α 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 $\tau \in \{p, ..., T\}$ as a function of the underlying short-run coefficient vector k (whose estimates are $\sqrt{T}$-consistent), of the super-consistent co-integration coefficients β, 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_τ:

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

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 τ. In the appendix we present the analytical form of this Jacobian.[5] 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_τ is given by:*

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

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_τ on the estimates is either non-existent (if a dummy variable for period τ was used) or asymptotically negligible. A standard system 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 τ) 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 τ and not as confidence bands for the entire

[5] 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.

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

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

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 forecasting. 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,

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

where the part ψ_{1t} represents the error-correcting movements of the system and is identical to the GG transitory component above, while the part ψ_{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 $p > 1$: $B^*(L) = B_0^* + B_1^* L + ... + B_{p-2}^* L^{p-2}$, where $B_j^* = \sum_{i=j+1}^{p-1} B_i$. Then it can be represented as a distributed lag of the observable variables:

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

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

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

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

$$\begin{aligned}
\tilde{\psi}_{2t} &= \psi_{2t} - \mathsf{E}(\psi_{2t}) \\
&= -(I-P)Q^{-1}B^*(L)\left(\Delta y_t - \mu^*\right) \\
&= \left(-[I-P]\,Q^{-1}\right)\left(B_0^*, -B_0^*\mu^*, B_1^*, -B_1^*\mu^*, ..., B_{p-2}^*, -B_{p-2}^*\mu^*\right) \times \\
&\quad (\Delta y_t', 1, \Delta y_{t-1}', 1, \ldots, \Delta y_{t-p+2}', 1)' \qquad (14) \\
&= \left(-[I-P]\,Q^{-1}\right)\left(B_0^*, B_1^*, ..., B_{p-2}^*, -B^*(1)\mu^*\right) \times \\
&\quad (\Delta y_t', \Delta y_{t-1}', \ldots, \Delta y_{t-p+2}', 1)'
\end{aligned}$$

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

$$
\begin{aligned}
\tilde{\psi}_t &= \tilde{\psi}_{1t} + \tilde{\psi}_{2t} \\
&= \left(P, -[I-P]\, Q^{-1}\left[B_0^*, B_1^*, \ldots, B_{p-2}^*\right], s_\mu\right) \times \qquad (15) \\
&\quad (y_t', \Delta y_t', \Delta y_{t-1}', \ldots, \Delta y_{t-p+2}', 1)' \\
&\equiv S(y_t', \Delta y_t', \Delta y_{t-1}', \ldots, \Delta y_{t-p+2}', 1)',
\end{aligned}
$$

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

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

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 α 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 coefficients, 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 $\tau \in \{p, \ldots, T\}$ as a function of k, of the co-integration coefficients β, and of the data; the only difference now is that we have to condition on the lagged values as well, $y_\tau, \ldots, y_{\tau-p+1}$:

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

The function f_{SW} is given by (15), $f_{SW} = S(y_t', \Delta y_t', \Delta y_{t-1}', \ldots, \Delta y_{t-p+2}', 1)'$. 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:*

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

Again the result follows directly from applying the delta method, cf. the remarks on Proposition 1, where now the conditioning data are given by $y_\tau, ..., y_{\tau-p+1}$. 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 τ 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 τ 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 $t = p...T$ by drawing from a suitable distribution describing the innovation process ε_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 $t = 0..p-1$. 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 γ 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

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

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.[6] 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

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

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

[6] In order not to overload the notation, we do not formally distinguish here between the true transitory component (true of course conditional on period-τ data) and its original point estimate, because we hope it is clear from the context that only the estimate can be used here.

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 Simulation study

In order to assess the relative empirical strengths and weaknesses of the delta method and the bootstrap variants we conduct a simulation study. The three chosen data-generating processes (DGPs) are bivariate, have a lag order of $p = 2$ and a single cointegration relationship, $\beta' = (1, -1)$. Both variables are reacting to equilibrium deviations, but with different speeds, $\alpha' = (-0.5, 0.25)$. The unrestricted constant term is chosen as $\mu' = (0.1, -0.01)$.

The three DGPs differ in their short-run dynamics. The first one has short-run coefficients that induce only small stationary roots. In particular, this "small root" DGP has

$$B_1 = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.2 \end{pmatrix},$$

which entails the roots 1, 0.5, $0.23 \pm 0.39i$ for the system polynomial. The second DGP instead uses

$$B_1 = \begin{pmatrix} 0.9 & 0.9 \\ 0.2 & 0.3 \end{pmatrix},$$

which gives the system roots 1, 0.91, $0.27 \pm 0.16i$, i.e. the largest stationary root is quite close to unity which is known to be troublesome in small samples. Finally, the third DGP is constructed to display the "common cycle" feature as discussed by Proietti (1997), i.e. the relationship between the transitory movements of the variables reflects the long-run adjustments, such that the (true) GG and SW components coincide. This is achieved by setting $B_1 = \alpha M$, where M is chosen as $M = (0.5, 0.3)$. The implied roots are 1, 0.46, -0.38, 0. However, we never treat the common-cycle feature as known at estimation time, i.e. we estimate B_1 freely.

We consider two different sample sizes for the simulation, $T = 100$ and $T = 300$. We simulate each case 2000 times drawing the innovations from a standard normal distribution, and for the bootstrap in each simulation run we use 1000 replications. As described before, the bootstrap resamples repeatedly from the respective estimated residuals. Of course simulating a bootstrap is computationally quite intensive.

As a further point of analysis, we distinguish between a cointegration relationship which is treated as known and thus fixed at its true coefficients β in each simulation run and in each bootstrap replication, and on the other hand estimating the cointegrating coefficients freely as $\hat{\beta} = (1, \hat{\beta}_2)$ and relying on their superconsistency. The latter case corresponds to the approach we choose in our empirical application in Section 6. Altogether we thus have three DGPs (small root, large root, or common cycle), two sample sizes (100 and 300), two estimation strategies (β known or not), two transitory components (GG and SW), two bootstrap variants (direct and Hall-type), and finally one analytical delta method. We consider the standard nominal test sizes of 1%, 5%, and 10%.

In this simulation we focus on the current edge of the sample. We report in Tables 1 and 2 the empirical rejection frequencies, i.e. the relative number of occurrences that the constructed confidence interval for the respective transitory component (GG or SW) did not include the true transitory component realized at the final observation for the given data constellation in each simulation run.

[Table 1 about here]

[Table 2 about here]

The first main simulation result is that the Hall-type bootstrap displays satisfactory size properties even for the case of large stationary roots, if the sample is at least moderately large (cf. Table 2). Therefore we recommend to use the Hall-type bootstrap in general and interpret rejections with some caution in small samples in the presence of large stationary roots. The second important –albeit not necessarily surprising– finding is that we observe some dramatic size distortions when the asymptotic delta method is used in small samples (Table 1), especially if the short-run dynamics are persistent (large-root DGP).

In between these good and bad news we have a differentiated picture where the direct bootstrap achieves improvements of the empirical coverage frequencies compared to the delta method, but it is still considerably oversized sometimes, while in some other cases it overcorrects the size distortion and yields a conservative test. The delta method is adequate especially when the cointegration coefficients are fixed, because by comparing Tables 1 and 2 we see that the convergence towards the nominal size in the case of unknown β is quite slow, and thus while the $\hat{\beta}$ estimate can be treated as fixed asymptotically, relying on this fact can be misleading for practical applications. Comparing the GG and the SW delta simulation results for β unknown it appears that the rejection frequencies are closer to the nominal level for the SW component. This is not surprising since the cointegration coefficients enter the additional component ψ_{2t} only indirectly and thus their nuisance variation is relatively less important for SW. Finally, for the common-cycle DGP, where the true GG and SW components coincide, we observe the expected effect that the results are similar for the GG and SW estimates.

6 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 use the quarterly variables real consumption $cons_t$, real (gross) investment inv_t, and the real "private" output yp_t obtained by subtracting real government expenditures from real GDP (all in logarithms). Instead of their sample 1947-1988 we analyze more recent data spanning 1974q1-2015q4 ($T = 168$). We also work with one common stochastic trend, letting the series be tied together by two co-integrating relationships ($r = 2$), such that any pair of two of the three variables is co-integrated. KPSW propose to specify the co-integrating relationships according to economic theory as the "great ratios" of balanced growth, specifically $cons - yp$ and $inv - yp$, but for the purposes of this illustration we will work with freely estimated co-integration coefficients β, see below for the estimates. 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 ($yp_t^{trans} = yp_t - yp_t^{perma}$, the negative of the output gap such that a negative sign corresponds to an actual gap, i.e. a shortfall of current output); Figure 1 shows the point estimates of the transitory components of both decompositions, GG and SW. Both estimates are quite similar for this data which may suggest the presence of common cyclical features (Proietti, 1997). The great recession of 2008/2009 is clearly visible as a large drop in excess output. In general we note that a quickly falling excess output estimate corresponds quite well to the NBER dating of the US recessions.

[Figure 1 about here]

As the error-correction terms $\beta' y_t$ are central ingredients for the transitory components, we display their (non-demeaned) estimates in Figure 2, associated with the cointegration coefficient estimates of $\hat{\beta} = [(1, 0, -0.97)', (0, 1, -1.17)']$. While the consumption-output vector is almost exactly a 1:1 relation, the investment-output great ratio would not be so great in terms of its stationarity properties after the publication of KPSW. A formal likelihood ratio test of the restrictions that impose the great ratios on the co-integration relations yields a rejection with $\chi^2(2) = 9.8$ and a p-value of 0.007. During the great recession the drop in (private) output was more pronounced than that of consumption, leading to an increase of the first error-correction term. In contrast, the second error-correction term fell dramatically due to an even larger drop of (private) investment. Comparing Figures 1 and 2 it is clear that the transitory output components mostly inherit their dynamics from the second error-correction term.

[Figure 2 about here]

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 α for the output equation. In our illustration here, this test yields the following result: $P(\chi^2(2) > 8.93) = 0.012$, 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 α 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 3. 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 is negligible compared to the rest of the sample. For the bootstraps we re-sample from the estimated residuals.

[Figure 3 about here]

The estimation uncertainty is considerable such that the point-wise output gap is not significantly different from zero for many observations, even for this relatively loose significance level. Nevertheless, for the three years from late 2008 through late 2011 the confidence intervals do not contain zero. This novel assessment of the estimation uncertainty therefore provides evidence that the output drop in the great recession did not (fully) stem from a movement of the permanent output component. Such a statement would not have been possible in a statistically rigorous manner by merely considering the point estimate.

As expected for theoretical reasons and from the simulation study, the delta method intervals tend to be tighter than their bootstrapped counterparts. The Hall-type bootstrap intervals are mostly similar to the direct bootstrap intervals, except in the year 2009 where they reach farther into the negative range. (The delta method intervals are of course symmetric around the respective point estimates of the gap.) Given the unknown cointegration coefficients and the medium-sized sample we favor the Hall-type bootstrap bands, following the insights from the simulation exercise in Section 5.

Finally, Figure 4 displays the corresponding measures and calculations for the SW decomposition of the output series in the co-integrated system. The comparison of the three different interval types yields similar patterns as with the GG decomposition: They include zero most of the times but not from early 2009 through mid 2011, and the delta method intervals are tighter. However, here the Hall bootstrap is even more similar to the direct bootstrap than in the GG case, so the implicit bias correction of the Hall approach appears relatively less important.

[Figure 4 about here]

7 Summary

While a permanent-transitory decomposition of non-stationary time series in a co-integrated 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 (say, τ) 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 period τ (possibly the latest available observation period). 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, SW) and Gonzalo-Granger (common-factor-based, GG) decompositions.

In our presentation we have focussed on the estimation uncertainty of the transitory components, but clearly the results can be directly applied to the permanent components as well, given the additivity of the decomposition $y_\tau = y_\tau^{perma} + y_\tau^{trans}$ (where y_t^{trans} is taken to have unconditional expectation zero). For some applications like output gap estimates it will be more natural to consider the transitory component, whereas in other cases it may be more interesting to draw the confidence intervals around the permanent component, for example when interpreting the permanent components of non-stationary unemployment rates as estimates of time-varying natural rates.

In the empirical illustration we calculated the uncertainty of output gap estimates for the US. For example, it turned out that for observations in the range from 2009 to 2011 the GG-based and SW-based gap estimates relating to private non-government output have 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 $(n, (r+(p-1)n+1))$:

$$Param = [\underset{(n,r)}{\alpha}, \underset{(n,n)}{B_1}, ..., B_{p-1}, \underset{(n,1)}{\mu}],$$

and we define the parameter vector as

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

A.1 Useful rules

For convenience we reproduce some of the differentiation rules and other useful formulas from LÃŒtkepohl (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):

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

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

$$vec(yx') = x \otimes y$$

- 10.4.1 (3), p.183:

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

- 10.6.3 (1), p.200:

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

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_τ, where $G(k;\beta) = Q^{-1}\alpha[\beta'Q^{-1}\alpha]^{-1}\left(\beta', [\beta'Q^{-1}\alpha]^{-1}\beta'Q^{-1}\mu\right)$ 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.:

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

We first formulate some of the parameters as explicit matrix products. For α we have

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

and thus:

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

Next, B_i is given as:

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

which implies:

$$\begin{aligned} vec(B_i) &\overset{7.2(5)}{=} \left(\left[0_{(n,r)},...,I_n,...,0_{(n,1)}\right]\otimes I_n\right)k \\ &= \left[0_{(n^2,nr)},0_{(n^2,n^2(i-1)},I_{n^2},0_{(n^2,n^2(p-1-i))},0_{(n^2,n)}\right]k \end{aligned} \tag{22}$$

In the same way, the estimated constant term μ can be expressed as:

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

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

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

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

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

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 $\left(0_{(r,n)},\tilde{\mu}\right) = \left(0_{(1,n)},1\right)\otimes\tilde{\mu}$, and after stacking we get:

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

So altogether we can use

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

or in stacked form:

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

Now we can calculate some derivatives of smaller parts:

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

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

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

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

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

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

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

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

$$
\frac{\partial vec(Q^{-1})}{\partial k'} \overset{10.6.3(1)}{=} -\left(Q'^{-1} \otimes Q^{-1}\right) \frac{\partial vec(Q)}{\partial k'} \qquad (30)
$$

Now the derivative of (the vector) μ is quite straightforward from (23):

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

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

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

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 $U \hat{=} Q^{-1}, V \hat{=} \alpha, Z^{-1} \hat{=} [\beta' Q^{-1}\alpha]^{-1}$, 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_τ enter. Thus:

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

where the embedded derivatives $\frac{\partial vec(Q^{-1})}{\partial k'}$, $\frac{\partial vec(\alpha)}{\partial k'}$, and $\frac{\partial vec(\beta' Q^{-1}\alpha)}{\partial k'}$ 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:

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

The embedded derivatives $\frac{\partial vec(\beta' Q^{-1}\alpha)}{\partial k'}, \frac{\partial vec(Q)}{\partial k'}, \frac{\partial \mu}{\partial k'}$ 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 τ only the matrix multiplications and additions in (33) have to be calculated that involve the conditioning data y_τ.

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 $\tilde{\psi}_{2,\tau} = \psi_{2,\tau} - \mathsf{E}(\psi_{2,\tau})$, we can separately analyze $\partial \psi_{2,\tau} / \partial k'$ and $\partial \mathsf{E}(\psi_{2,\tau}) / \partial k'$, and subtract them afterwards.

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

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

where the $0_{(1,j)} \otimes I_{n^2}$ element comes from the lacking $B_1 ... B_j$ 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 $B^*(L) = B_0^* + B_1^*L + ... + B_{p-2}^*L^{p-2}$, and as before $P = Q^{-1}\alpha(\beta'Q^{-1}\alpha)^{-1}\beta'$. This can also be written with explicit data vectors:

$$\psi_{2,\tau} = -(I-P)Q^{-1}\left[B_0^*, ..., B_{p-2}^*\right]\begin{bmatrix} \Delta y_\tau \\ \Delta y_{\tau-1} \\ \vdots \\ \Delta y_{\tau-(p-2)} \end{bmatrix}$$

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

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

The derivative in the second line (dealing with Q) is known from equ. (29). The derivative in the first line is equal to $\partial vec(P)/\partial k' = \partial vec\left(Q^{-1}\alpha(\beta'Q^{-1}\alpha)^{-1}\beta'\right)/\partial k'$, which is of course very similar to the result in equ. (33), except that $(\beta', \tilde{\mu})(y_\tau', 1)'$

is replaced by just β' at the end. Therefore:

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

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

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

where in the last line we have exploited the formula 2.4(13) for column vectors. Now we can find the derivative of $vec\left[B_0^*,\ldots,B_{p-2}^*\right]$ by applying formula 10.4.4(8):

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

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

$$\frac{\partial vec\left[B_0^*,\ldots,B_{p-2}^*\right]}{\partial k'} = \sum_{j=0}^{p-2}\left(\left[\begin{array}{c}0_{(j,1)}\\1\\0_{(p-2-j,1)}\end{array}\right]\otimes I_{n^2}\right)\times \tag{38}$$
$$\left[0_{(n^2,nr)},0_{(1,j)}\otimes I_{n^2},1_{(1,p-1-j)}\otimes I_{n^2},0_{(n^2,n)}\right]$$

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 $\partial\mathsf{E}(\psi_{2,\tau})/\partial k'$

We still need to consider the de-meaning term:

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

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

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

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

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 ($U\hat{=}B^*(1)$ and $V\hat{=}I_n - P$, the rest are identity matrices), we get:

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

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

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

where equ. (35) has been used.

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

References

CHANG, Y., B. JIANG, AND J. Y. PARK (2012): "Using Kalman Filter to Extract and Test for Common Stochastic Trends," Discussion paper.

CHANG, Y., J. I. MILLER, AND J. Y. PARK (2009): "Extracting a common stochastic trend: Theory with some applications," *Journal of Econometrics*, 150(2), 231–247.

Cottrell, A., and R. Lucchetti (2013): *Gretl User's Guide* Version 1.9.14; 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 Economics 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.

LÃŒtkepohl, H. (1997): *Handbook of Matrices*. Wiley.

Morley, J. C. (2002): "A state-space approach to calculating the Beveridge-Nelson decomposition," *Economics Letters*, 75, 123–127.

Oh, K. H., E. Zivot, and D. Creal (2008): "The relationship between the Beveridge-Nelson decomposition and other permanant-transitory decompositions that are popular in economics," *Journal of Econometrics*, 146(2), 207–219.

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.

B Tables

Table 1: Simulation results, small sample

$T=100$		β unknown				β known			
		GG		SW		GG		SW	
		y_1	y_2	y_1	y_2	y_1	y_2	y_1	y_2
	nominal	*DGP with small stationary root*							
Delta	1	10.3	1.9	6.9	3.0	2.9	1.0	3.5	3.2
	5	18.9	6.0	14.1	7.6	8.6	4.5	9.0	8.4
	10	25.8	12.0	20.2	13.1	14.6	8.3	14.5	13.1
Bootstrap direct	1	1.4	0.0	1.1	0.9	1.5	0.0	1.1	0.9
	5	6.5	0.1	6.4	4.2	6.6	0.1	6.3	4.1
	10	12.2	0.7	12.6	9.3	12.2	0.6	11.7	9.5
Bootstrap Hall	1	0.9	1.6	1.0	1.1	1.1	2.0	1.1	1.2
	5	5.0	7.6	6.2	7.2	4.2	7.8	4.3	7.4
	10	9.5	14.8	11.7	13.2	9.0	15.3	9.9	13.8
		DGP with large stationary root							
Delta	1	19.3	22.0	26.8	25.7	19.0	22.2	26.0	25.1
	5	27.5	30.0	33.7	32.1	27.4	29.8	33.3	32.5
	10	33.6	34.7	38.2	37.5	32.7	34.8	37.6	37.6
Bootstrap direct	1	6.0	7.4	9.9	9.8	5.8	6.7	10.2	9.7
	5	18.7	21.8	26.2	25.0	17.6	20.6	25.4	25.2
	10	29.3	33.8	38.1	36.8	28.7	33.8	37.1	37.3
Bootstrap Hall	1	1.1	4.3	4.2	3.1	1.1	4.5	5.0	3.2
	5	9.4	15.9	14.7	13.0	8.0	13.5	14.9	13.6
	10	19.0	28.7	28.0	25.6	18.4	26.1	27.9	26.2
		Common-cycle DGP							
Delta	1	9.9	5.6	9.5	4.3	3.6	1.5	3.2	1.2
	5	18.7	12.9	16.4	11.0	8.1	5.6	7.2	4.2
	10	25.0	19.4	21.6	17.3	12.7	10.7	11.4	9.1
Bootstrap direct	1	2.3	0.7	1.7	0.7	1.9	0.7	1.6	0.7
	5	7.2	4.9	6.4	4.4	7.8	3.9	6.7	3.7
	10	13.9	10.1	12.3	9.3	13.0	10.4	11.9	9.5
Bootstrap Hall	1	1.1	2.0	1.4	1.6	0.7	1.6	0.7	1.6
	5	5.3	6.8	6.3	7.6	4.7	6.9	5.6	6.7
	10	10.5	12.4	11.1	13.8	10.3	11.9	11.4	11.9

Given numbers are relative rejection frequencies in per cent. 2000 simulation runs, for the bootstrap 1000 replications in each run. For the definition of the DGPs see the text.

Table 2: Simulation results, larger sample

$T = 300$		β unknown				β known			
		GG		SW		GG		SW	
		y_1	y_2	y_1	y_2	y_1	y_2	y_1	y_2
	nominal	*DGP with small stationary root*							
Delta	1	7.5	1.8	4.9	1.3	1.8	0.9	1.5	1.6
	5	15.6	7.0	11.3	6.0	5.6	4.5	5.5	5.6
	10	23.4	12.8	17.6	10.1	11.0	9.6	10.6	9.7
Bootstrap direct	1	1.1	0.0	1.3	0.7	1.4	0.1	1.1	0.7
	5	5.5	0.5	5.7	4.5	5.6	1.1	5.4	5.3
	10	10.7	3.5	11.9	9.3	11.0	4.2	10.7	10.2
Bootstrap Hall	1	0.6	1.4	0.8	1.1	1.1	1.7	1.0	1.3
	5	4.7	7.2	4.9	6.0	4.9	6.8	4.5	5.8
	10	9.3	14.2	10.4	12.0	9.4	13.7	10.4	11.0
		DGP with large stationary root							
Delta	1	8.1	9.7	10.5	9.5	6.8	8.7	9.5	8.8
	5	13.7	15.8	16.6	15.5	12.4	14.0	15.9	15.4
	10	19.3	20.5	21.3	20.8	17.9	18.8	21.0	20.5
Bootstrap direct	1	2.3	3.3	2.5	2.8	2.0	3.0	3.0	2.7
	5	9.9	11.5	11.1	11.0	8.2	10.9	10.3	9.6
	10	16.6	19.0	19.8	18.6	15.8	19.4	18.2	17.3
Bootstrap Hall	1	0.4	1.1	0.2	0.2	0.5	0.6	0.3	0.4
	5	4.8	5.2	4.6	4.5	4.2	4.7	3.4	3.4
	10	9.7	11.8	10.9	10.5	10.2	11.4	9.3	8.7
		Common-cycle DGP							
Delta	1	7.4	4.7	6.1	3.0	1.3	1.5	1.8	1.1
	5	15.4	12.0	14.0	9.4	5.1	5.6	6.0	4.6
	10	21.8	17.6	20.7	16.8	10.1	10.9	10.6	10.0
Bootstrap direct	1	1.1	0.9	1.1	0.7	1.2	1.3	1.0	1.3
	5	5.7	4.8	5.4	4.2	5.8	5.6	5.8	5.4
	10	10.8	10.5	10.2	9.3	11	11.2	10.3	10.6
Bootstrap Hall	1	0.7	1.1	0.8	1.0	1.4	1.0	1.4	1.0
	5	4.2	4.7	5.1	4.9	4.9	4.9	4.9	5.3
	10	8.8	9.1	9.6	9.7	10.4	8.9	9.6	9.7

Given numbers are relative rejection frequencies in per cent. 2000 simulation runs, for the bootstrap 1000 replications in each run. For the definition of the DGPs see the text.

C Figures

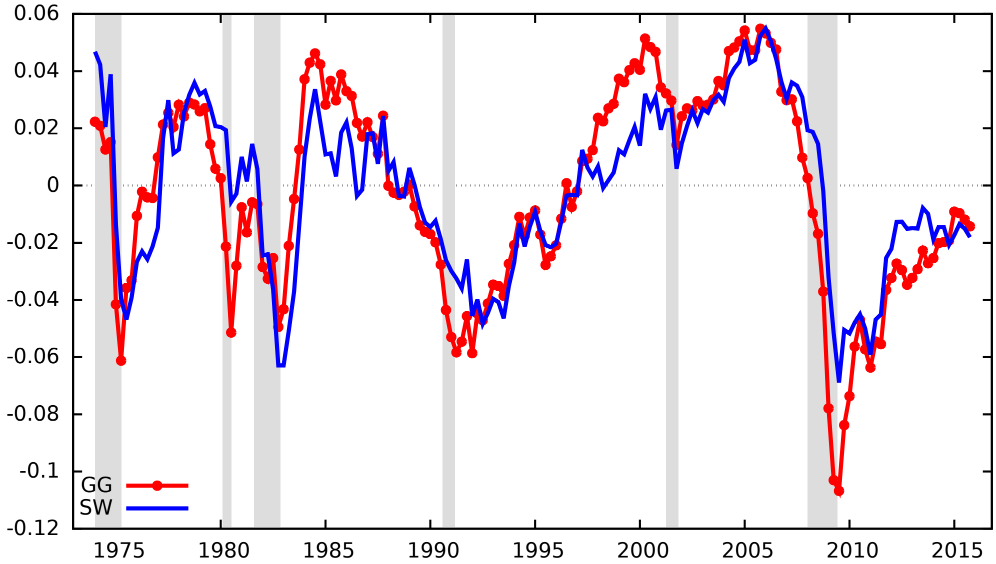


Figure 1: Estimated output gaps as transitory components of the GG and SW permanent-transitory decompositions. Shaded areas indicate NBER recession dating.

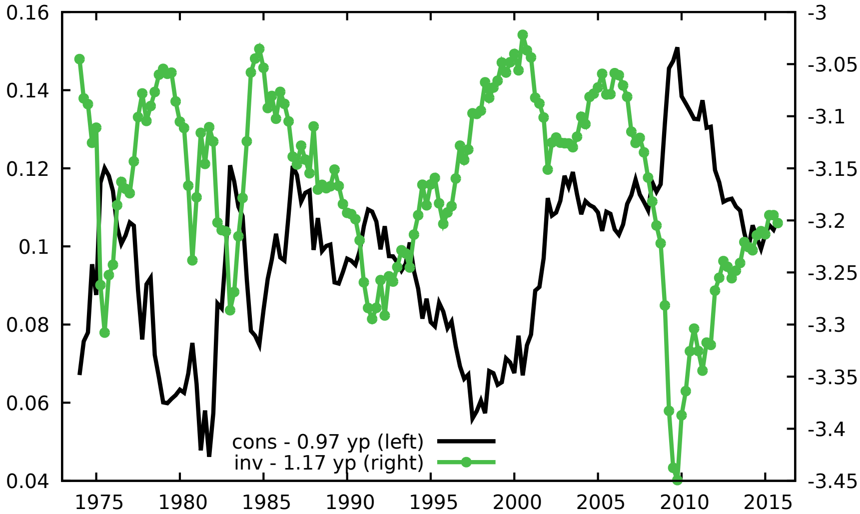


Figure 2: Error correction terms. These are directly taken as $\hat{\beta}' y_t$ without subtracting their expected value. Also note the different ranges.

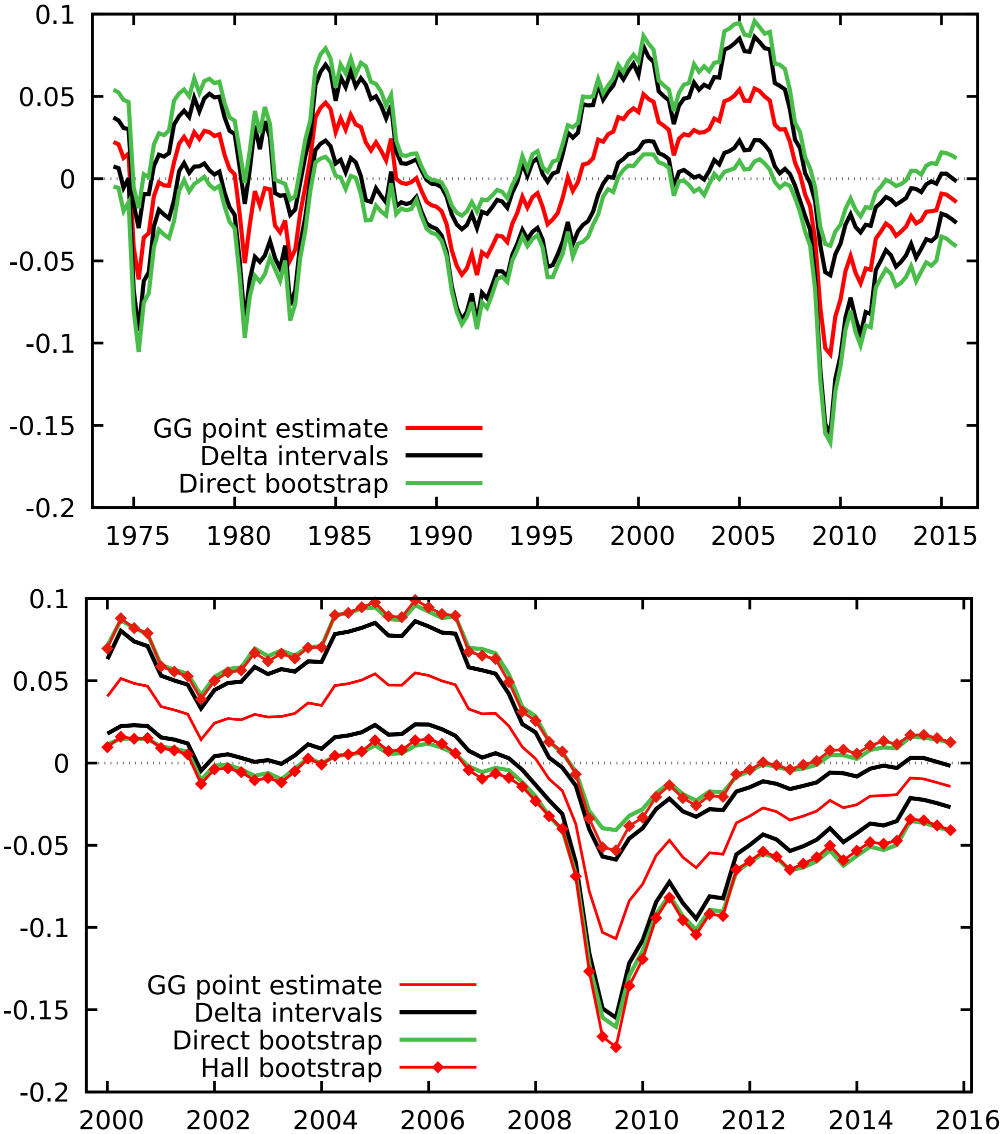
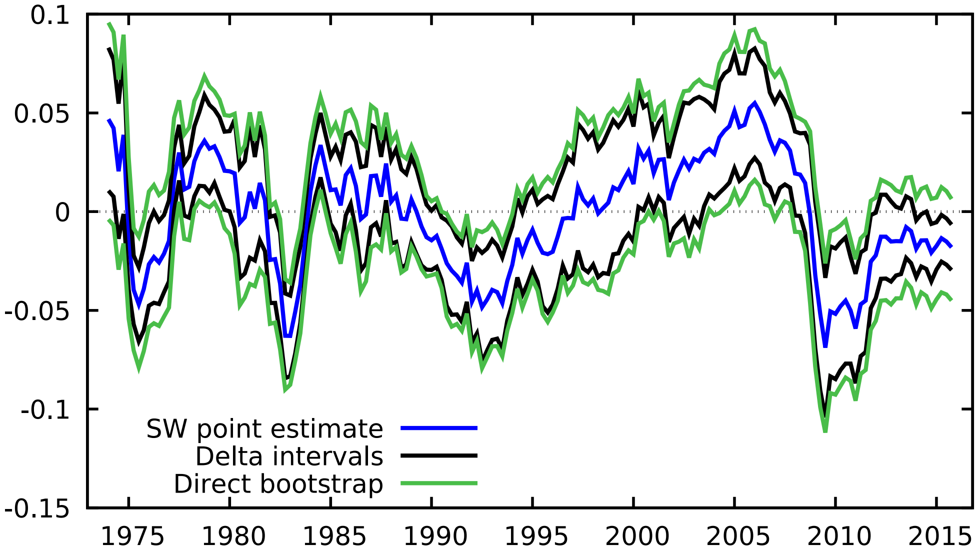


Figure 3: GG decomposition, confidence intervals for the output gap; displayed together for all periods in the sample for convenience, while the interpretation should be for a single period only.
Upper panel: all periods, without Hall bootstrap results to avoid clutter. Lower panel: calculation as before, but zooming in on the latter part of the sample, and with Hall bootstrap.



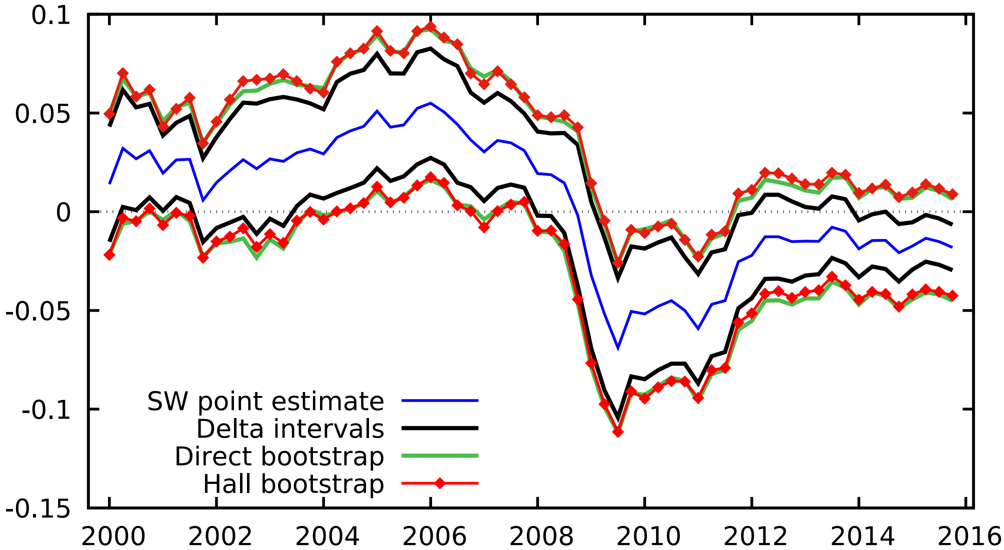


Figure 4: SW decomposition, confidence intervals for the output gap; displayed together for all periods in the sample for convenience, while the interpretation should be for a single period only.
Upper panel: all periods, without Hall bootstrap results to avoid clutter. Lower panel: calculation as before, but zooming in on the latter part of the sample, and with Hall bootstrap.