Discussion Paper No. 08-065

A Critical Note on the Forecast Error Variance Decomposition

Atılım Seymen

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 08-065

A Critical Note on the Forecast Error Variance Decomposition

Atılım Seymen

Non-Technical Summary

This paper discusses the limitations and applicability of the forecast error variance decomposition (FEVD) in business cycle analysis. Upon closer inspection, a lacking connection between FEVD and the usual definition of the business cycle is found, which in turn favours the use of a historical variance decomposition (HVD) of variables for assessing the driving forces of cyclical fluctuations. The business cycle definition implied by FEVD is shown to be particularly problematic if the underlying vector autoregressive (VAR) process comprises nonstationary variables, since FEVD would then imply a nonstationary cyclical component. This, however, contradicts the widespread consensus among macroeconomists about the stationary nature of cycles regardless of whether the underlying series for the measurement of the cycle is nonstationary.

In our analysis, the FEVD technique is scrutinized and arguments against using it in the business cycle context are brought forward. Then, as an alternative to FEVD, the HVD approach is presented and discussed. Empirical applications in the paper employ both the Hodrick-Prescott (HP) filter and the Beveridge-Nelson-decomposition (BND) to assess the reasonability of FEVD in the business cycle context and to compare it with the HVD. The results show the spurious nature of the conclusions yielded by the former approach. In particular, through confronting historical forecast errors of output at different forecast horizons based on two different empirical models with its cyclical component computed with the HP-filter and the BND, it is found that the FEVD analysis based on the particular empirical models leads to spurious conclusions regarding the nature of business cycles.

All in all, the historical variance decomposition approach is shown to overcome the problems related to the FEVD, since it has the advantage of being directly compatible with conventional business cycle definitions and hence, being indirectly able to solve the nonstationarity problem related to the FEVD. In addition to producing more intelligible results, the HVD approach is shown to cover all the statistical properties of times series referred to in a business cycle analysis.

Nicht-technische Zusammenfassung

Dieses Papier diskutiert die Grenzen und Anwendungsmöglichkeiten der Prognosefehlervarianzzerlegung (FEVD) für Konjunkturanalysen. Bei genauerer Betrachtung wird der mangelnde Zusammenhang zwischen der FEVD und den üblichen Definitionen von Konjunkturzyklen ersichtlich, sodass als Konsequenz die Anwendung der historischen Varianzzerlegung (HVD) favorisiert wird. Die durch FEVD implizierte Definition von Konjunkturzyklen ist besonders problematisch, falls der zugrundeliegende vektorautoregressive (VAR) Prozess nicht-stationäre Variablen beinhaltet, da die FEVD in diesem Fall zu einer nicht-stationären zyklischen Komponente führt. Das widerspricht jedoch dem unter Makroökonomen verbreiteten Konsens über die stationäre Natur von Zyklen, unabhängig davon ob die für die Bemessung der Zyklen zugrunde liegende Zeitreihe nicht-stationär ist.

In unserer Analyse untersuchen wir die FEVD-Methode und bringen Argumente gegen seine Anwendung im Zusammenhang mit Konjunkturzyklen an. Darüber hinaus wird die HVD-Methode als eine Alternative zur FEVD dargestellt und erörtert. Um die Vernünftigkeit der FEVD-Methode einschätzen und sie mit der HVD-Methode vergleichen zu können, werden in empirischen Anwendungen sowohl der Hodrick-Prescott (HP) Filter als auch die Beveridge-Nelson-Zerlegung (BND) eingesetzt. Die Resultate zeigen die Fragwürdigkeit der von der FEVD-Methode hervorgebrachten Ergebnisse auf. Insbesondere durch Vergleich historischer Prognosefehler, die für verschiedene Prognosehorizonte berechnet werden und auf zwei unterschiedlichen empirischen Modellen basieren, wird ersichtlich, dass die FEVD-Analyse zu Scheinergebnissen im Hinblick auf die Eigenschaften der Konjunkturzyklen kommt.

Zusammenfassend lässt sich zeigen, dass der HVD-Ansatz die Probleme, die durch den FEVD-Ansatz entstehen, bewältigt, da er den Vorteil hat, direkt mit konventionellen Definitionen von Konjunkturzyklen überein zustimmen und demzufolge indirekt dazu fähig ist, das Nicht-Stationaritätsproblem, das sich durch die FEVD-Methode ergibt, zu lösen. Die HVD-Methode führt nicht nur zu Ergebnissen, die eindeutig interpretierbar sind, sondern nimmt auch Bezug auf alle statistischen Merkmale von Zeitreihen, die in einer Konjunkturanalyse in Betracht kommen.

A Critical Note on the Forecast Error Variance Decomposition

Atılım Seymen*

August 2008

Abstract

The paper questions the reasonability of using forecast error variance decompositions for assessing the role of different structural shocks in business cycle fluctuations. It is shown that the forecast error variance decomposition is related to a dubious definition of the business cycle. A historical variance decomposition approach is proposed to overcome the problems related to the forecast error variance decomposition.

JEL classification: C32, E32

Keywords: Business Cycles, Structural Vector Autoregression Models, Forecast Error Variance Decomposition, Historical Variance Decomposition

---

*I am greately indebted to Bernd Lucke, Ulrich Fritsche and Olaf Posch for their comments and suggestions on earlier drafts of the paper. The work also benefited from participation in the following: a seminar at the University of Hamburg, 2nd Workshop on Macroeconomics and Business Cycles (Makroökonomik und Konjunktur) at the ifo Institute for Economic Research in Dresden, the EC2 Conference 2007 on Advances in Time Series Econometrics in Faro, at the XIII. Spring Meeting of Young Economists in Lille. Finally, I received a great deal of support while writing large parts of the paper during my tenure as a research assistant at the Institute for Growth and Business Cycles of the University of Hamburg. Any errors are my own.

Center for European Economic Research (ZEW), P.O. Box 103443, D-68304 Mannheim, Germany. E-mail: seymen@zew.de

1. Introduction

Forecast error variance decomposition (FEVD) is an econometric tool used by many economists in the vector autoregression (VAR) context for assessing the driving forces of business cycles. Given that many macroeconomic models can also be written in the VAR form, FEVD can be applied in empirical as well as theoretical models. This study shows that the connection between FEVD and the popular definitions of the cycle[2] is not well established. Therefore, it advocates employing a historical variance decomposition (HVD) of variables for business cycle analysis. The latter decomposition is shown to provide a remedy to the problems related to FEVD.

Several issues arise when economists want to investigate the sources of cyclical fluctuations, the first of which is the identification of structural shocks and their dynamic effects, both of which are not directly observable. The structural VAR (SVAR) literature offers a multitude of possibilities for identifying - for example - supply, demand, technology, monetary policy, etc. shocks and the dynamic response of macroeconomic variables to them. Alternatively, the parameters of a theoretical model can be calibrated / estimated and the recursive law of motion of its variables can be written in VAR form. In both types of models, the shocks and their dynamic multipliers determine the properties of the cyclical fluctuations. Yet the cycle itself is still not directly observable, which is the second issue to be dealt with. The most common techniques used by applied economists for extracting the cyclical component of the data include differencing, computing two-sided moving averages, filtering out linear / quadratic time trends, and filtering in frequency domain.[3] Finally, a method is needed for determining the contribution of different structural shocks to the cycle. FEVD is one of such methods, which is claimed to be non-suitable for business cycle analysis in this paper.

The main argument of this paper claims that FEVD is related to a dubious business cycle definition, which renders trend and cycle components of variables that are quite different than the ones suggested by conventional business cycle definitions. The business cycle definition implied by FEVD is particularly problematic if the underlying VAR process comprises

[2]We henceforth use the terms "business cycle", "cycle" and "cyclical fluctuation" interchangeably.
[3]See Baxter and King (1995) for a review of these techniques.

nonstationary variables, since FEVD would then imply a nonstationary cyclical component. There is consensus among macroeconomists, on the other hand, that cycles are stationary, while many macroeconomic time series are nonstationary.[4] Moreover, the typical usage of FEVD implies a one-to-one relationship between forecast errors and cyclical components, which is far from being even approximately true.

In order to present our point of view, we employ the filter proposed by Hodrick and Prescott (1997; henceforth HP-filter) and the Beveridge-Nelson-decomposition (BND) proposed by Beveridge and Nelson (1981) in the applications in this paper, which are widely implemented approaches for measuring business cycles. Two empirical findings of this paper strengthen the case against FEVD in the business cycle context. The first finding comes from the confrontation of the historical forecast errors of output at different forecast horizons with its cyclical component computed with the HP-filter and the BND based on the models by Gali (1999) and King et al. (1991; henceforth KPSW). We find that historical forecast errors at lower frequencies (i.e. longer forecast horizons) are as highly correlated with business cycles as within the so-called business cycle horizon. Second, historical forecast errors of output are found to be nonstationary according to unit root tests when the forecast horizon exceeds 10 quarters in both Gali and KPSW models. This finding implies that a FEVD analysis based on these popular models leads to spurious conclusions for a forecast horizon longer than ten quarters.

Finally, even if the forecast errors were related one-to-one to business cycle fluctuations, it should not be forgotten that the business cycle is typically defined to be a macroeconomic phenomenon that occurs in a certain time span, say 6 to 32 quarters. FEVD does, however, not deliver which macroeconomic shocks are the main driving force of the busines cycle fluctuations over the *entire* business cycle horizon. Historical variance decomposition provides, on the other hand, information for the entire business cycle horizon, whatever its length is assumed to be according to the chosen business cycle measure.

Note that this paper does *not* discuss which business cycle definition is the most appropriate one to be used by macroeconomists, but emphasizes merely that FEVD, a commonly followed econometric approach for business cycle analysis in time series models, is not con-

[4]See Baxter and King (1995) on widely accepted time-series properties of business cycles.

sistent with widely used definitions of the cycle. HVD, which is based on the idea that historical cycles of macroeconomic variables can be decomposed with respect to (w.r.t.) structural shocks, is shown, on the other hand, to be compatible with different definitions of the cycle.

The outline of the paper is as follows. The next section presents the FEVD technique and the arguments against using it in the business cycle context. Section 3 illustrates the HVD approach. Section 4 gives examples of the implementation of the HVD. Concluding remarks are provided in Section 5.

2. Forecast Error Variance Decomposition

2.1. Stationary VAR

Forecast error variance decomposition is carried out typically based on the moving average (MA) representation of a (stationary) $VAR(p)$ process with p being the order of the VAR,

$$X_t = CD_t + \sum_{i=0}^{\infty} \Theta_i w_{t-i}, \tag{1}$$

where X_t is a $K \times 1$ vector of endogenous variables, Θ_i is the i^{th} $K \times K$ MA coefficient matrix, w_t is a $K \times 1$ vector of orthogonal white noise innovations all with a unit variance, C is an $(K \times M)$ coefficient matrix corresponding to the deterministic terms represented by the $(M \times 1)$ matrix D_t.[5] One can write the h-step forecast error for the process as

$$X_{t+h} - X_t\left(h\right) = \sum_{i=0}^{h-1} \Theta_i w_{t+h-i}, \tag{2}$$

with $X_t\left(h\right)$ being the optimal h-step forecast at period t for X_{t+h}. It is straightforward to compute the total forecast error variance of a variable in X_t for the h-step forecast horizon and the corresponding shares of individual innovations to this variance, see Lütkepohl (2005). What is traditionally done in the literature is to set h such that the computation is made for the business cycle horizon.[6]

[5] M is here the number of deterministic variables such as constant, time trend, dummies, etc.

[6] This means setting $6 \leq h \leq 32$ if you work with quarterly data following the business cycle definition used by many macroeconomists.

In the structural VAR literature the covariance matrix of the structural innovations w_t is typically restricted to be an identity matrix without loss of generality. When decomposing the forecast error variances, it is furthermore assumed that the structural innovations do not exhibit any autocorrelation and correlation among their leads/lags. The total forecast error variance of the variables in X_t are given, under the aforementioned restrictions and assumptions by the diagonal elements of

$$E\left[\left(X_{t+h}-X_t\left(h\right)\right)\left(X_{t+h}-X_t\left(h\right)\right)'\right]=\sum_{i=0}^{h-1}\Theta_i\Sigma_w\Theta_i'=\sum_{i=0}^{h-1}\Theta_i\Theta_i', \tag{3}$$

where Σ_w is the covariance matrix of the structural innovations and is set to be the K-dimensional identity matrix. The contribution of the k^{th} structural shock to the forecast error variance of the j^{th} variable for a given forecast horizon is computed by

$$\sum_{i=0}^{h-1}\left(e_j'\Theta_i e_k\right)^2 \tag{4}$$

where e_k is the k^{th} column of the K-order identity matrix. Given (3) and (4), it is straightforward to compute the share of a structural shock in the fluctuations of a variable.

2.2. Nonstationary VAR

Equation (1) represents a VAR with stationary variables. However, since the influential work of Nelson and Plosser (1982), many macroeconomic variables are known/assumed to be nonstationary due to a unit root. Therefore, this study also focuses on models with nonstationary variables. Regardless of whether the process includes cointegrated variables, every nonstationary VAR possesses an MA representation in first differences given by

$$\Delta X_t=\mu+\sum_{i=0}^{\infty}\Theta_i w_{t-i}, \tag{5}$$

where Δ is the difference operator such that $\Delta X_t = X_t - X_{t-1}$, and μ is a $K\times 1$ constant vector.[7] Hence, it can be easily shown that the total forecast error variance of the variables in X_t are given by the diagonal elements of the matrix $\sum_{i=0}^{h-1}\Theta_i^*\Theta_i^{*\prime}$ with $\Theta_i^*=\sum_{j=0}^{i}\Theta_j$ for a

[7]Thus, the only deterministic term here is chosen to be a constant for the sake of presentation. This implies a linear time trend for the process which is observed in many macroeconomic time series. Moreover, X_t is assumed to consist of $I\left(1\right)$ variables, which have only one unit root so that first-differencing renders them stationary.

forecast horizon of h. The contribution of the k^{th} structural shock to the forecast error variance of the j^{th} variable for a given forecast horizon is analogously given by $\sum_{i=0}^{h-1}\left(e_j'\Theta_i^* e_k\right)^2$.

Equation (5) can be rewritten as

$$X_t = X_0 + \mu t + \Theta_0^* w_t + \Theta_1^* w_{t-1} + \cdots + \Theta_{t-1}^* w_1 \tag{6}$$

with X_0 being the vector containing the initial value of the process. Note that (6) gives an exact representation of the realisation of X_t, $t = 1, \ldots, T$ with estimated Θ_i^*, $i = 0, 1, \ldots$ and w_i, $i = 1, \ldots, T$, where T stands for the number of observations in the sample excluding the initial values.

2.3. FEVD and Measurement of Cycles

This subsection elaborates the connection between forecast error variance decomposition and typical definitions of the business cycle. We first briefly review two different approaches to business cycle measurement that are widely used in the empirical literature and will be employed in the empirical applications of this paper. Then, we present the case against using FEVD in the business cycle context.

2.3.1. Two Popular Measures of Cycles

The HP-filter introduced by Hodrick and Prescott (1997) has been widely used in the empirical literature when investigating the properties of cycles in the last two decades. As will be shown in the following, it implies that macroeconomic processes have a non-linear and non-deterministic trend. In order to see this, notice that X_t has three components according to (6): a (stochastic) constant (X_0), a linear time trend (μt) and a further stochastic component $\left(\Theta_0^* w_t + \Theta_1^* w_{t-1} + \cdots + \Theta_{t-1}^* w_1\right)$. Since the HP-filter is a linear filter, it is applied to all of these three components when computing the cyclical component of X_t. It removes the constant and the linear time trend entirely from X_t and the stochastic component only partly. Therefore, according to the HP-filter, structural innovations contribute not only to the cycles of the variables in X_t, but to their long-run trends as well.

The trend-cycle decomposition proposed by Beveridge and Nelson (1981) underlies a different philosophy than the HP-filter and is applicable only to nonstationary processes.

The trend component of X_t is a random walk with drift according to the BND and reads

$$\tau_t = X_0 + \mu t + \Theta(1) \sum_{i=1}^{t} w_i \tag{7}$$

with τ_t denoting the trend component and $\Theta(1) = \Theta_0 + \Theta_1 + \ldots$ the matrix of long-run multipliers for the model in (5).[8] Hence, the BND also implies that structural innovations contribute to the long-run trend of X_t.

2.3.2. The Case against FEVD in the Business Cycle Context

The trend and cyclical components implied by forecast error variance decomposition are quite different from the ones implied by the HP-filter and the BND. For a sample which starts at period 1 when the initial values needed for the VAR estimation are excluded, FEVD implies a cyclical term of the form

$$\Theta_0^* w_t + \Theta_1^* w_{t-1} + \cdots + \Theta_{h-1}^* w_{t-h+1},\ t \geq h \tag{8}$$

and a trend term of the form

$$X_0 + \mu t,\ t = h \text{ and } X_0 + \mu t + \Theta_h^* w_{t-h} + \cdots + \Theta_{t-1}^* w_1,\ t > h. \tag{9}$$

Notice that the sum of the trend and cyclical component from (8) and (9) gives the total represented by (6).[9] (8) and (9) imply that the properties of the cycle, and therefore of the long-run component, differ w.r.t. h under FEVD.[10] However, the stability of the VAR in (5) implies that Θ_i^* and Θ_j^* are approximately equal for i and j big enough. Hence, the expression in (9) for $t > h$ can be written as

$$X_0 + \mu t + \Theta_h^* w_{t-h} + \cdots + \Theta_{t-1}^* w_1 \approx X_0 + \mu t + \Theta^* (w_{t-h} + \cdots + w_1) \tag{10}$$

since $\Theta^* \approx \Theta_h^* \approx \Theta_{h-1}^* \approx \cdots \approx \Theta_{t-1}^*$ when h is large enough. In such a case, FEVD implies that cyclical fluctuations occur around a long-run trend consisting of a linear time trend and a multivariate random walk. After which value of h (10) holds almost perfectly is a question

[8] Beveridge and Nelson (1981) consider a univariate model, but the implementation in the multivariate case is straightforward.

[9] For a stationary VAR model of the form (1), the cyclical and trend components are analogously given by $\Theta_0 w_h + \Theta_1 w_{t-1} + \cdots + \Theta_{h-1} w_{t-h+1}$, $t \geq h$ and CD_t, $t = h$ and $CD_t + \Theta_h w_{t-h} + \cdots + \Theta_{t-1} w_1$, $t > h$.

[10] See the discussion below on the relationship between business cycle horizon and FEVD.

that we turn to in Section 2.4. Note that this definition of trend can be rather different than the one following from HP-filtering or Beveridge-Nelson decomposition.

Figure 1 shows the trend and cyclical components of the output data used by Gali (1999) computed with a trend based on FEVD for a forecast horizon of 12, the HP-filter and the BND.[11] Estimated trend and cyclical components of a variable are obviously very sensitive to how they are measured, as the reported features of the cyclical components in Table 1 shows. Those differ quite a bit w.r.t. their volatility, amplitude, persistence and comovement properties.

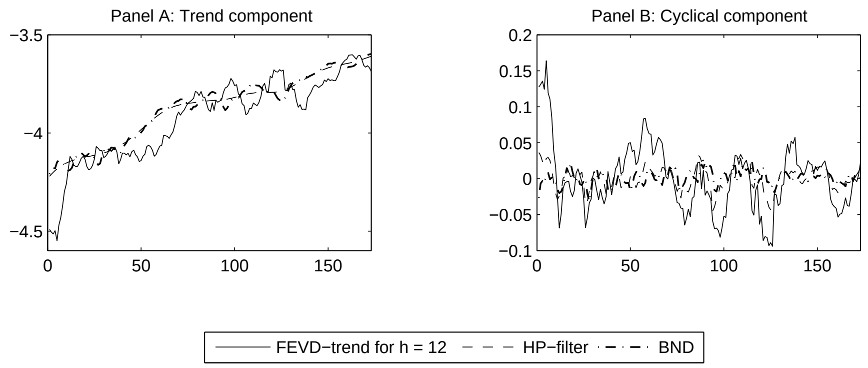

Figure 1: Cyclical and trend components of output according to FEVD-trend for $h = 12$, HP-filter and Beveridge-Nelson-decomposition in Gali model

In this regard, another issue when using the FEVD in the business cycle context is the ambiguity about why setting h to a value within the business cycle horizon should render a business cycle analysis. The estimated cross-correlation coefficients between the historical i-horizon forecast errors, computed based on both Gali (1999) and KPSW models, and the cyclical component of output are displayed in Figure 2. A closer relationship between the historical forecast errors at business cycle frequencies and the business cycle fluctuations of output cannot be established; that is, the correlations are not particularly higher at the so-called business cycle frequencies (i.e., at a forecast horizon of six to thirty-two quarters) than at lower frequencies.

[11]Note that the cycles computed with the FEVD have a non-zero mean, while the HP-cycles and the BND cycles do have a zero mean. We have normalized the trend following from the FEVD and the cycles around it accordingly by subtracting the mean from the cyclical component and adding it to the corresponding trend component.

Table 1: Characteristics of cyclical components according to Gali model

	Relative volatility	Amplitude	Persistence
FEVD-Cycles	1.00	0.26	0.90
HP-Cycles	0.35	0.10	0.85
BN-Cycles	0.21	0.07	0.63

	Correlation	
	FEVD-Cycles	HP-Cycles
HP-Cycles	0.71	
BN-Cycles	0.05	0.31

Notes: Relative volatility is the standard deviation of one series divided by the standard deviation of the FEVD-cycles for a forecast horizon of 12. Amplitude stands for the difference between the maximum and minimum values of the cycles shown in Figure 1. The persistence measure is the estimated coefficient of the first lag of the corresponding cycle in an AR(1) model.

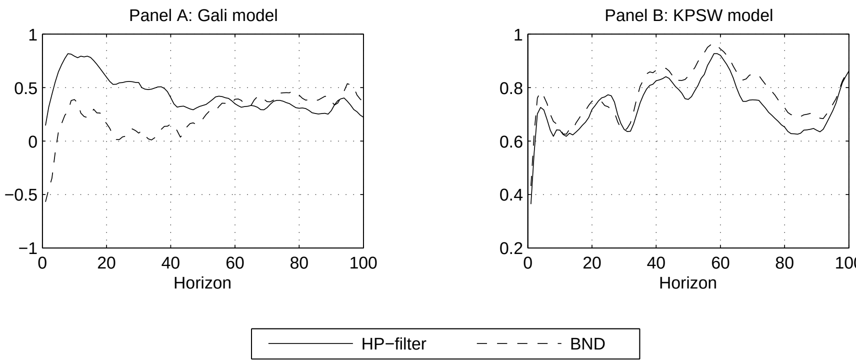

Figure 2: Cross-correlation between i-horizon historical forecast errors and cyclical component of output according to Gali and KPSW models

As a final remark in this context, abstracting from the problem that the connection between business cycles and forecast errors is rather ambigious, the FEVD technique does not give an answer to the question of which structural shocks drive the business cycle fluctuations over the *entire* business cycle horizon. The result obtained by KPSW, who identify "balanced-growth", "inflation" and "real-interest-rate" shocks, provides a good example. It is reported in their study that the fraction of the forecast error variance of output attributable to the real-interest-rate shock is 74 percent in 4 quarters, 55 percent in 8 quarters and 25 percent in 24 quarters. Moreover, the contribution of the technology shock to output fluctuations is forecasted to be 5 percent in 4 quarters, 22 percent in 8 quarters and 62 percent in 24 quarters. KPSW conclude that technology and real-interest-rate shocks are both important in the cyclical fluctuations. However, given that a typical definition of the business cycle horizon is from 6 to 32 quarters, it is not possible to have an idea about the weight of real-interest-rate and technology shocks in the fluctuations of output over the *entire* business cycle horizon. Whether the share of real-interest-rate shocks is 55 percent, 22 percent, or something in between, for example, cannot be assessed. Similarly, it is not clear whether technology shocks play a bigger role than real-interest-rate shocks over the entire business cycle horizon.

2.4. Nonstationarity and FEVD

Since many macroeconomic time series are known to be nonstationary, this sub-section deals with the consequences of applying the FEVD based on a VAR with nonstationary variables. It is shown in the following that a statistical problem arises when the underlying SVAR model, such as the one represented by equation (5), comprises nonstationary variables. In that case, the cyclical component of the variables implied by FEVD, i.e. the forecast errors, become nonstationary when the forecast horizon approaches infinity. Therefore, forecast error variance decomposition leads merely to spurious conclusions after a certain forecast horizon. We are interested in this paper whether this critical forecast horizon is already reached within the domain of the so-called business cycle frequencies.

In order to illustrate our point of view, we rewrite the h-step forecast error for a nonstationary VAR given by (8) as

$$z_t^h = \theta^* (L) w_t, \tag{11}$$

where z_t^h stands for the h-step forecast error and $\theta^* (L) = \Theta_0^* + \Theta_1^* L + \cdots + \Theta_{h-1}^* L^{h-1}$. Obviously, as h approaches infinity, $\theta^* (L)$ approaches $\Theta^* (L)$. Notice that $\Theta^* (L)$ governs the motion of X_t, which is a nonstationary process by construction. Hence, it can be concluded that z_t^h approaches nonstationarity with increasing h. In order to get an intuition for this result, note that the h-step forecast error can be approximated by

$$z_t^h \approx \Theta_0^* w_t + \Theta_1^* w_{t-1} + \cdots + \Theta_k^* w_{t-k} + \Theta^* (w_{t-k-1} + \cdots + w_{t-h+1}) \tag{12}$$

with increasing h, see Equation (10) in the previous sub-section. z_t^h approaches nonstationarity with increasing h due to two effects. First, the coefficient matrices converge, that is, $\Theta^* \approx \Theta_h^* \approx \Theta_{h-1}^* \approx \cdots \approx \Theta_{t-1}^*$ becomes a better approximation when h is larger. Second, the term $w_t^h := w_{t-k-1} + \cdots + w_{t-h+1}$ increasingly yields a random walk character with higher h. Note that w_t^h can be summarised by $w_t^h = w_{t-1}^h + w_{t-k-1} - w_{t-h}$. Obviously, more w_{t+j}^h share a common component for $j < h$ with increasing h.[12] Hence, since the multiplication of a constant matrix with a multivariate random walk series also renders a random walk series and, moreover, sum of stationary and nonstationary series renders a nonstationary series, forecast errors should also be approximately nonstationary for a large enough h.

In practice, it is important to decide what the critical forecast horizon k is, i.e. after which forecast horizon the "cyclical term" $\Theta^* w_t^h$ is highly persistent and therefore very close to nonstationarity. We check the convergence properties of the dynamic multipliers of output represented by Θ_i^* in the Gali model as an example. It can be seen in Figure 3 that after roughly the 10$^{\text{th}}$ to 15$^{\text{th}}$ coefficient matrix the dynamic multipliers of output with respect to technology and nontechnology shocks converge to their long-run values. Furthermore, w_t^h for $h = 4, 12$ for the technology shocks estimated by the same model are shown in Figure 4. The increasing random walk character of the series with increasing h is immediately clear from the figure. Augmented-Dickey-Fuller (ADF) test also confirms this intuition by rejecting nonstationarity of w_t^{12}, but not of w_t^4.

[12] A numerical example may make this point more clear. Let $t = 12$ and $h = 12$. Moreover, let $k = 0$ without loss of generality for the result that we would like to show. Then, w_{t+j}^{12} for $j = 1, \ldots 12$ all contain w_{12}^{12}.

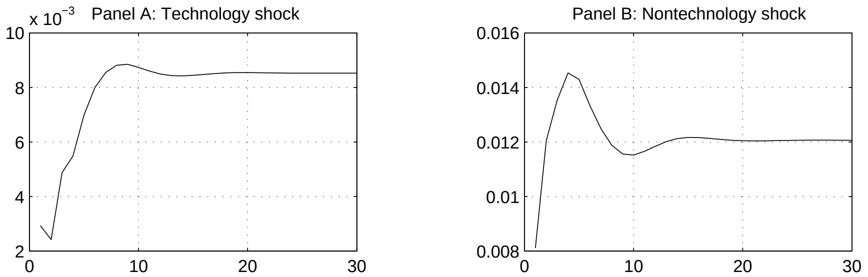

Figure 3: Dynamic multipliers of output with respect to technology and nontechnology shocks in the Gali model

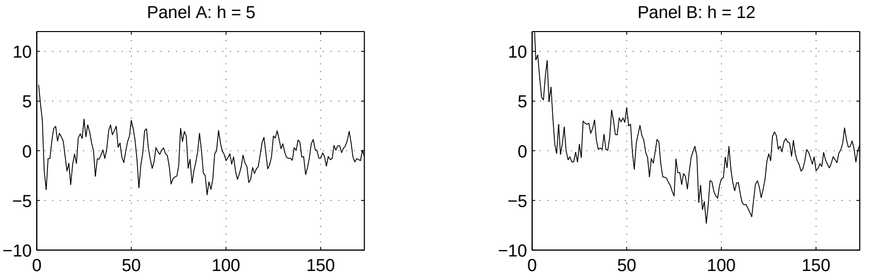

Figure 4: Accumulated technology shocks according to the Gali model within a chosen horizon

Also, the critical forecast horizon for output forecast errors based on Gali and KPSW models were checked by directly looking at the unit root properties of the estimated historical forecast error series. Although the two models have different structures and different definitions of the output variable, the historical forecast error series of output becomes nonstationary after a forecast horizon of ten according to the ADF test in both models. It can be accordingly concluded that it is questionable to carry out a FEVD analysis of output in these models, with the given data set, for a forecast horizon of $h > 10$ since the reported error variance shares is very likely to be spurious.

3. Historical Variance Decomposition

Modern (theoretical) business cycle models, for example dynamic stochastic general equilibrium (DSGE) models, are designed such that unforeseeable shocks lead to cyclical fluctuations through macroeconomic propagation mechanisms. Given that such shocks are at the same time the source of forecast uncertainty, macroeconometrists employ a decomposition of the variance of forecast errors when they want to establish the driving forces of business cycle fluctuations. Yet, the previous section has shown that the connection between conventional business cycles and forecast errors is rather ambigious and, therefore, FEVD is not a suitable tool for business cycle analysis. Furthermore, FEVD has been shown to imply a nonstationary business cycle component of output, which contradicts with the idea of recurrent cyclical fluctuations, since stationarity is a necessary property for the recurrence of cycles.

In this section, we discuss a different type of variance decomposition that we call historical variance decomposition (HVD). HVD can be employed in the SVAR context based on conventional business cycle definitions. Thus, the first problem related to using the FEVD in the business cycle context is resolved directly and the second problem indirectly, since conventional business cycle definitions imply stationary cyclical fluctuations. HVD is conducted when the cyclical component of a variable, say x_t, can be decomposed into its sub-components with respect to the structural shocks, w^i $i = 1, \dots N$, where N is the number of structural shocks in the VAR and $K \geq N$:

$$x_t = x_{t,w^1} + \cdots + x_{t,w^N} \tag{13}$$

x_{t,w^i} being the realisation of x_t had only the i^{th} structural shock occured from the beginning of the sample until period t.[13]

Note that (13) holds exactly in the case of a linear model, like an SVAR model, and when the cyclical component is computed with a linear filter, like the HP-filter. In order to see this, let $\tilde{x}_t$ be one of the variables in a nonstationary process of the form (6), which is given by

$$\tilde{x}_t = \tilde{x}_{t,w^1} + \cdots + \tilde{x}_{t,w^N}, \tag{14}$$

where $\tilde{x}_{t,w^i}$ is analogous to x_{t,w^i} and the related term following from $X_0 + \mu t$ is excluded. If x_t is defined as the cyclical component of $\tilde{x}_t$ computed with a certain linear filter, then x_{t,w^i} is the cyclical component of $\tilde{x}_{t,w^i}$ for $i = 1, \dots N$ computed with the same filter. Therefore,

[13] In many cases, $K = N$. One example for where this is not the case is the bivariate model of Gali (1999) with three variables and two structural shocks.

the cyclical component of an HP-filtered variable which is the sum of multiple additive components is, for example, equal to the sum of the cycles of its HP-filtered sub-components.

A natural path to follow for assessing the role of different structural shocks in cyclical fluctuations is to compute estimates of their shares in the variance of x_t. One possible estimate stems from the approximation

$$var\left(x_t\right) \approx var\left(x_{t,w^1}\right) + \cdots + var\left(x_{t,w^N}\right), \tag{15}$$

where $var\left(x\right)$ stands for the variance of the variable x. $var\left(x_{t,w^i}\right)/var\left(x_t\right)$ provides then an estimate of the share of the i^{th} structural shock in the variance of x_t. (15) holds only approximately in practice, since the zero correlation among w_t and w_{t-i} condition for $i \neq 0$ is typically only assumed, but not imposed by SVAR identification schemes. Therefore, $var\left(x_{t,w^i}\right)$ $i = 1, \ldots N$ do usually not add up to one in practice.

An alternative to (15) could be the statistical identity

$$var\left(x_t\right) = cov\left(x_t, x_{t,w^1}\right) + \cdots + cov\left(x_t, x_{t,w^N}\right), \tag{16}$$

where $cov\left(x, y\right)$ stands for the covariance between the variables x and y. In this case, $cov\left(x_t, x_{t,w^i}\right)/var\left(x_t\right)$ is the estimate of the share of the i^{th} structural shock in the variance of x_t. Since (16) represents an identity and hence holds exactly, the estimated shares add up to one. However, it should be noted that this estimation is ad hoc and does not have a theoretical justification. What is done with a decomposition like (16) is merely distributing the empirically observed, but unwanted nonzero covariances among w_t and w_{t-i} for $i \neq 0$ evenly and adding them to $var\left(x_{t,w^i}\right)$. An important remark in this context is that a covariance must not necessarily be positive. In practice, a negative $cov\left(x_t, x_{t,w^i}\right)$ must be relatively very small in absolute value and should then follow from a very low $var\left(x_{t,w^i}\right)$ and some negative $cov\left(x_{t,w^i}, x_{t,w^j}\right)$ for $i \neq j$. If, however, $cov\left(x_t, x_{t,w^i}\right)$ is so big such that, for example, the estimated share $cov\left(x_t, x_{t,w^i}\right)/var\left(x_t\right)$ bears a negative value smaller than -0.05, this should be seen as the consequence of a misspecified model. It may be, for example, that not enough lags have been included in the estimated VAR so that the residuals show autocorrelation. Another possibility is that the sample data is subject to some extreme events and includes some outliers, but those have not been accounted for with dummys or some exogenous variables.

A historical variance decomposition of the form (15) or (16) is always possible whenever the cyclical term can be decomposed as in (13). Therefore, it can be shown that the Beveridge-Nelson decomposition is also compatible with HVD. The cyclical component of X_t is given by $\sum_{i=0}^{t-1} \Psi_{i+1} w_{t-i}$ with $\Psi_j = \sum_{i=j}^{\infty} \Theta_{i+1}$ and $w_{1-i} = 0$ for $i > 0$, which is a linear expression and can therefore be easily rewritten in the form of (13).

An important aspect of variance decomposition, be it FEVD or HVD, analysis is that business cycles are typically characterised by properties like volatility, amplitude, persistence and co-movement, and a variance decomposition analysis refers to all of these. In order to see this on the example of HVD, note that $var\left(x_{t,w^1}\right) \approx cov\left(x_t, x_{t,w^1}\right)$ are approximately equal when the model is well specified, see the equations (15) and (16) and the discussion above. Moreover, $cov\left(x_t, x_{t,w^1}\right)$ is given by

$$cov\left(x_t, x_{t,w^i}\right) = \rho\left(x_t, x_{t,w^i}\right) \sqrt{var\left(x_t\right)} \sqrt{var\left(x_{t,w^1}\right)}, \tag{17}$$

by definition, where $\rho\left(x, y\right)$ stands for the correlation between x and y. The expression in (17) refers to the comovement between a cyclical term and one of its sub-components through the correlation coefficient as well as volatility, amplitude, persistence through the standard deviation terms given by $\sqrt{var\left(x\right)}$ for the variable x. Hence, the share of a shock in the variance of x_t depends on all these properties.

To summarise, historical variance decomposition resolves all the problems related to forecast error variance decomposition discussed in the previous section. It is compatible with a multitude of business cycle definitions. Since those automatically imply a stationary cyclical component, even when the underlying series is nonstationary, a meaningful variance decomposition analysis is yielded. Furthermore, the business cycle analysis refers to the entire business cycle horizon, however long it may be defined by the business cycle measure used, and not to only a certain time point within the business cycle horizon. Finally, generally accepted characteristics of cycles are all referred to by such an approach.

4. Empirical Applications

Two empirical applications of the HVD approach are presented in this section. The first empirical application follows from the bivariate model of Gali (1999), which comprises the labor productivity and the total hours worked. Gali motivates and employs an identification scheme for estimating technology and nontechnology shocks mentioned before in this paper. An important property of his VAR model is that output is also indirectly included in the estimation, since that variable is just the sum of labor productivity and total hours worked by definition. Hence, conducting any analysis of output dynamics like for labor productivity and total hours worked is an accessible and easy task.

Gali (1999) is interested in whether technology or nontechnology shocks drive business cycles. In order to assess it, he first establishes the high correlation between the cyclical components of output and hours worked according to the HP-filter, and then shows that nontechnology shocks lead to a strong positive comovement between the two variables according to this SVAR model, but not technology shocks. Therefore, his conclusion is that cyclical fluctuations are driven by nontechnology shocks. A HVD based on cycles measured with the HP-filter implies the same conclusion by attributing a share of 0.82 and 0.86 to nontechnology shocks in the variance of output cycle according to the formulas in (15) and (16), respectively, see Table 2.

When the cycle is defined, however, according to the Beveridge-Nelson decomposition in the Gali model, technology shocks turn out to be the driving force of output fluctuations with a share of 0.67/0.69 according to Table 2. This result already implies that the conclusion Gali (1999) arrives at in his study has a lot to do with the business cycle definition used by him. Note that when his analysis is based on the BND, the cyclical components of output and hours are still very highly correlated with a coefficient of 0.87. Yet the strong positive comovement does not follow only from a strong comovement of nontechnology components in this case, but the technology components are highly correlated as well. The former comovement corresponds to a correlation coefficient of 0.96 and the latter to a coefficient of 0.95. The reason behind the larger share of technology shocks in the BN-cycles is that the technology BN-component of output has a variance, which is 2.81 times larger than the variance of its nontechnology component.

KPSW estimate only structural shocks with permanent effects and their identification scheme differs from Gali's such that the model they work with embodies cointegrating relationships. There are three identified structural shocks, mentioned above already, and three

Table 2: Shares of structural shocks in the variance of output fluctuations according to historical variance decomposition in the Gali model

	Technology	Nontechnology
FEVD-Cycles	0.26	0.74
HP-Cycles	0.11	0.82
	[0.14]	[0.86]
Beveridge-Nelson-Cycles	0.67	0.29
	[0.69]	[0.31]

Notes: Two estimates of the shares are provided for HP-cycles and Beveridge-Nelson-Cycles. The first reported share is computed based on (15), while the second estimate (in squared brackets) is based on (16). FEVD estimates of shares are provided for a forecast horizon of 12.

transitory shocks in their model, which are not attributed an economic interpretation. FEVD implies for this model that balanced-growth and real-interest-rate shocks have almost equally the highest share in the variance of forecast errors with $h = 12$, see Table 3. This finding can, however, not be confirmed by HVD based on both the HP-filter and the BND, according to which the real-interest-rate shocks must have been alone the most important driving force of output cycles. Finally, transitory shocks have a share of merely 0.15 according to the FEVD with $h = 12$, while HVD with both business cycle definitions attributes them a share of about 0.30.

An important issue is that the estimated shares with (15) and (16) do not differ much for neither the Gali nor the KPSW model. The largest difference is observed for the estimated share of nontechnology shocks in the Gali model with 0.04. As discussed in the previous section, differences bigger than 0.05 should rather be taken as an indication of model misspecification and the empirical model must be corrected accordingly.

5. Concluding Remarks

A commonly investigated research topic of modern macroeconomics is the driving forces of cyclical fluctuations. The most important challenge for macroeconomists is that cycles and the shocks, which are the driving forces of cycles, are not directly observable. When the driving forces of business cycles is investigated, three core questions need to be answered: i)

Table 3: Shares of structural shocks in the variance of output fluctuations according to historical variance decomposition in the KPSW model

	Balanced-growth	Inflation	Real-interest-rate	Transitory
FEVD-Cycles	0.44	0.03	0.39	0.15
HP-Cycles	0.11	0.04	0.57	0.28
	[0.09]	[0.06]	[0.56]	[0.29]
Beveridge-Nelson-Cycles	0.24	0.03	0.45	0.27
	[0.21]	[0.06]	[0.44]	[0.29]

Note: See Table 2.

How should the structural shocks and their dynamic effects be identified? ii) How should the business cycle be defined? iii) How should the contribution of structural shocks to the cycle be computed? The focus of this paper has been the third question criticising the forecast error variance decomposition technique as a tool of business cycle analysis.

It has been shown that the FEVD is related to a business cycle definition, which is quite different then the conventional business cycle definitions used in the macroeconomic literature. Furthermore, it has been shown to produce spurious results when applied in models with nonstationary variables. A historical variance decomposition approach has been claimed to overcome the problems related to the FEVD, since it has the advantage of being directly compatible with conventional business cycle definitions and hence, of being indirectly able to solve the nonstationarity problem related to the FEVD. Moreover, the HVD has been shown, like the FEVD, to cover all the statistical properties of time series referred to in a business cycle analysis.

The HVD has been implemented based on both the HP-filter and the BND for illustrating the amount of distortion that the FEVD techniqe may cause in business cycle analysis. Note that the models by Gali (1999) and KPSW (1991) used in the implementations comprise different variables and use different identification approaches. They have been taken as they are and the estimatiton were carried out with the original data sets of the corresponding studies. The results with the HVD have been confronted then with the orinigal results of Gali (1999) and KPSW. Not so surprisingly, the results have shown that the business cycle definition underlying the HVD has a strong influence on the findings. Hence, whether the business cycle definition implied by the FEVD is an important question.

References

Baxter, M. and R. G. King, 1995, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *NBER Working Paper*, No. 5022

Beveridge, S. and C. R. Nelson, 1981, "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, Vol. 7, 151-174

Gali, J., 1999, "Technology, Employment and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *The American Economic Review*, Vol. 89, No. 1, 249-271

Hodrick, R.J. and E.C. Prescott, 1997, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, Vol. 29, No. 1, 1-16

King, R., Plosser, C., Stock, J. and M. Watson, 1991, "Stochastic Trends and Economic Fluctuations," *The American Economic Review,* Vol. 81, No. 4, 819-840

Lütkepohl, H., 2005, *New Introduction to Multiple Time Series Analysis,* Springer Verlag, Berlin

Nelson, C.R. and C.I. Plosser, 1982, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, Vol. 10, 139-162