

Priesmeier, Christoph

Conference Paper

Fiscal policy over the business cycle in Germany - a tale in three acts. Evidence from a time-varying VAR analysis 1970-2008

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Empirical Analyses of Fiscal Policy, No. B9-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Priesmeier, Christoph (2010) : Fiscal policy over the business cycle in Germany - a tale in three acts. Evidence from a time-varying VAR analysis 1970-2008, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Empirical Analyses of Fiscal Policy, No. B9-V2, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37206>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Fiscal policy over the business cycle in Germany - a tale in three acts

Evidence from a time-varying VAR analysis 1970-2008

Gerrit B. Koester Christoph Priesmeier*

Deutsche Bundesbank
Humboldt University Berlin

Verein für Socialpolitik
Jahrestagung 2010
Kiel - September 2010

Draft

JEL-Codes: E62, E32, D78

Abstract

Large parts of the empirical economic literature state that policy makers have ex-post proven to be unable to time fiscal policy measures countercyclically. In this paper we expand the usually applied empirical approaches and allow for changing reaction patterns over time by employing not only VAR- but as well time-varying parameter VAR estimation techniques on quarterly fiscal data for Germany from 1970-2008. Our analysis shows not a general a-cyclical pattern but three distinct phases of different fiscal policy timing indicating possible mis-specifications of time-invariant analyses.

*Views are explicitly those of the authors and not those of Deutsche Bundesbank.

Contents

1	Introduction	3
2	Timing of fiscal policy over the cycle - the literature	4
3	Measurement concepts and data-set	6
4	Evidence based on time-invariant VAR analysis	9
4.1	The benchmark model: The reduced form time-invariant parameter VAR	9
4.2	Identification of an output gap shock	10
4.3	Benchmark results	12
5	A time-varying parameter (TVP) VAR approach	14
5.1	The reduced form (homoskedastic) time-varying parameter VAR	14
5.2	The normal linear state space representation of the TVP-VAR	15
5.3	Identification of an output gap shock in the TVP-VAR	20
5.4	Estimation strategy: Maximum likelihood estimation of the state space model - A three stage estimation procedure	20
5.5	Initialization: Priors	22
5.5.1	Least squares priors	23
5.5.2	An alternative approach: Some literature priors	24
5.6	Estimation results	25
5.6.1	Maximum likelihood estimates	25
5.6.2	The optimal variances and covariances	25
5.7	Structural analysis	26
5.7.1	Analysis of cyclically unadjusted data in a time-varying frame- work:	26
5.7.2	Analysis of cyclically adjusted data in a time-varying framework:	29
6	Conclusion and outlook	31
7	Literature	32
8	Appendices	36
8.1	Appendix A:	36
8.1.1	Appendix A:i Stability Analysis for time-invariant benchmark model (Choleski identification): UNADJUSTED DATA	36
8.1.2	Appendix A:ii Stability Analysis for time-invariant benchmark model (Choleski identification): ADJUSTED DATA	38
8.1.3	Appendix A:iii Forecast-Error-Responses:	40

8.2	APPENDIX B:	40
8.2.1	B.1 The log-likelihood function	40
8.3	APPENDIX C:	41
8.3.1	APPENDIX C. 1. An alternative initialization approach	41
8.4	Appendix D: Estimation Results	44

List of Figures

1	Fiscal Policy over the business cycle in Germany 1970-2008	7
2	Impulse responses benchmark model	13
3	Time-varying parameter SVAR Impulse responses - cyclically unadjusted data	27
4	Time-varying parameter SVAR Impulse responses - cyclically unadjusted data	28
5	Time-varying parameter SVAR Impulse responses - cyclically unadjusted data	28
6	Time-varying parameter SVAR Impulse responses - cyclically adjusted data	29
7	Time-varying parameter SVAR Impulse responses - cyclically adjusted data	30
8	Time-varying parameter SVAR Impulse responses - cyclically adjusted data	31

1 Introduction

The estimation of fiscal multipliers has - since the publication of Keynes general theory in 1936 - not ceased to play an important role in the empirical economic literature. The massive economic stimulation packages, which have recently been implemented in nearly all leading industrial economies to cushion the effects of the financial crisis on the real economy, have even amplified the general interest in fiscal multipliers.

The timing of fiscal policy measures on the other hand has received far less attention and the empirical methods applied to analyze the timing have often been far less advanced than those in fiscal multiplier analyses. This is surprising as the effectiveness of fiscal policy in smoothing the business cycles depends crucially on the ability to implement fiscal measures in time.

Partly the limited attention to the timing of fiscal policy might be due to the general analytical problems of such analyses. First, changes in the fiscal stance and the business cycles development are likely to affect each other, which causes an endogeneity problem. Second, changes in the fiscal stance are not only caused by discretionary fiscal policy but are affected as well by the working of automatic stabilizers. Here the challenge is to separate discretionary from automatic responses to the business cycle. Third, the relationship between the business cycle and the fiscal stance might not be time-invariant. This poses problems e.g. for standard time-invariant time series analysis. Finally, a reliable empirical analysis needs to rely on a sufficiently large and long data-set of fiscal indicators, which is often not available.

With this paper we want to contribute to the empirical analysis of the timing of fiscal policy over the business cycle. To tackle the general endogeneity problem and to account for the fact that fiscal policy develops over time we propose modern Vector-Autoregressive (VAR) time-series analyses. We perform such analyses for unadjusted fiscal indicators as well as for a set of cyclically adjusted variables, which allows us to distinguish between the timing of discretionary fiscal policy and the working of automatic stabilizers. To account for possible regime changes in the timing of fiscal policy in different periods we include time-varying VAR analyses, expanding the empirical apparatus applied so far in the study of the timing of fiscal policy. The data-set we use build on the Deutsche Bundesbank national accounts data-base in quarterly frequency from 1970-2008, which gives us a large data-set with 156 observations - suitable for time-series analysis.

In our analyses we find a stable pattern of automatic stabilizers throughout the sample. The time-variant analyses show important changes in the timing of discretionary fiscal policy from 1970 to 2008. We find three distinct phases from 1970 to 1978, 1978 to 1990 and 1990 to 2000. While discretionary policy measures were timed

countercyclically in the first phase, the second phase showed a strong pro-cyclical timing. During the third phase - after German reunification - the timing of discretionary fiscal policy was improved again and measures were implemented slightly countercyclically. The three different phases did not coincide with changes in the political constellation of the federal government in 1982 and 1998, which indicates that the partisan orientation of government is unlikely to have been decisive for the general changes in the timing of fiscal policy. Our findings stand in contrast to large parts of the current empirical literature, which finds a general a-cyclical timing of fiscal policy measures over the business cycle - if analyzed from an ex-post perspective. Our results indicate that the received wisdom of an a-cyclical timing of fiscal policy is likely to result from empirical misspecifications, which we avoid by allowing for time-varying responses of fiscal policy to the business cycle.

The paper proceeds as follows: In part two we shortly review some important contributions in the literature and outline our general research question. Part three presents the data and the indicators employed and discusses the methods applied to transform and adjust the data. In part four we present a time-invariant benchmark analysis of the timing of fiscal policy, which we think is misspecified as there are important indicators pointing at parameter instability. To account for this finding we perform a time-variant analysis in part five, which points at strongly time-varying regimes of fiscal policy timing. The last part concludes.

2 Timing of fiscal policy over the cycle - the literature

In the literature we find multiple approaches to the study of the timing of fiscal policy over the business cycle. The first - and most common - approach is to distinguish discretionary fiscal policy from the working of automatic stabilizers by cyclically adjusting revenue and expenditure developments (see e.g. Alesina and Perotti 1995 or Giavazzi and Pagano 1996). In this branch of the literature cyclical adjustment usually takes place by subtracting a "cyclical" component from the aggregate revenue and expenditure developments. This cyclical component is calculated based on an indicator for the state of the business cycle (usually the output gap) and an elasticity measure for the effects of business cycle fluctuations on fiscal developments. The evidence from this literature points to an a-cyclical timing of fiscal policy - at least as far as only the deficit or the general fiscal stance is concerned. Gali and Perotti (2003) found that cyclically adjusted deficits have not reacted to the business cycle in Europe after the signing of the Maastricht treaty. These results are at least partly confirmed in other studies like Ballabriga and Martinez-Mongay (2002) or Wyplosz (2006). Separate analyses of cyclically adjusted spending and expenditure develop-

ments seem to indicate that although the fiscal balance might point at a-cyclicality of fiscal policy, discretionary revenue policies seem to be timed counter-cyclically while discretionary expenditure policies tend to have been pro-cyclical in Europe (see e.g. Fatas and Mihov (2009) or Turrini (2008)).

A second line of research emerged from the criticism, that fiscal adjustment processes are unable to unmask discretionary fiscal policy (see e.g. Chalk (2002) or Larch and Salto (2005) and for an early discussion Blanchard (1990)). One argument in this respect is that the elasticity of fiscal variables to the macroeconomic development might not be invariant over time (see e.g. Jaeger and Schuknecht (2004)) and might not cover all the relevant cyclically sensitive spending categories (see e.g. Darby and Melitz 2008). This line of research relies therefore on data of the estimated impact of legislative changes, which is e.g. collected by Romer and Romer (2007a,b, 2008) for the US. For the Euro area Agnello and Cimadomo (2009) have analyzed legislative changes for the revenue side and found that revenue policies seem to have been very pro-cyclical Europe in recent years, challenging the findings of Fatas and Mihov (2009) or Turrini (2008). However, there are several problems with this kind of analysis as well. Researchers usually have to rely on government forecasts of the expected fiscal impact of fiscal policy measures, which might be inaccurate and biased. Furthermore such estimates are usually only available for a limited number of countries and years. Additionally nearly all studies in this tradition have to focus on the revenue side, as comparable data on changes of spending are largely unavailable.

A third branch of research argues that the timing of fiscal policy does not only need to be analyzed *ex post*, but as well based on the information available at the time of passing the measures. This literature relies on real-time data about budget plans and expectations about the macroeconomic development. The number of studies is still relatively limited, but so far the results tend to indicate that fiscal plans show a stronger counter-cyclicality than *ex-post* data (see e.g. Beetsma and Giuliadori (2008), Giuliadori and Beetsma (2007) or Cimadomo (2007) for related studies).

In this paper we want to contribute to the literature on the timing of fiscal policy over the business cycle with a study of the developments in Germany from 1970-2008 on a quarterly basis. The main innovation of our approach is that we do not only estimate time-invariant relationships but employ a time-varying VAR as well. We apply this approach to cyclically adjusted as well as unadjusted data. By estimation and comparison of the results of time-varying and time-invariant approaches to adjusted and unadjusted data we want to contribute to the following questions:

1. Has the timing of fiscal policy over the cycle shown a stable pattern over the last 40 years in Germany or have there been different regimes at different times?
2. Have the results of time-invariant models been distorted and therefore unreliable

because of a time-variant pattern?

3. Do the results depend crucially on the employed method of cyclical adjustment of the government's fiscal stance?
4. Does the data indicate a stable or a time-variant working of automatic stabilizers

To keep the analyses - despite the application of time-invariant and time-variant models and adjusted and non-adjusted data - halfway parsimonious, we restrict ourselves here to an analysis of the government's fiscal stance (defined as general government revenues minus general expenditures excluding interest spending; both including social security and according to the national accounts definition) and the real output gap. However, an extension of the approach to separate analyses of expenditure and revenue developments and even a disaggregation of different spending and revenue components seems worthwhile in future research.

3 Measurement concepts and data-set

As our research focus lies on the timing of fiscal policy over the business cycle from an ex-post perspective, we need indicators for the fiscal policy stance and for the state of the business cycle.

With respect to the state of the business cycle we rely on the real GDP gap, which we have calculated based on the quarterly national accounts database of Deutsche Bundesbank. Nominal GDP was first realized by the chain-linked GDP deflator and then seasonally adjusted. In a second step we applied a HP-Filter ($\lambda=1600$) to the real GDP series which we prolonged with its own trend in the past (1960-1970) and the future (2009-2019) to avoid a distortion of the results at the lower and upper bound of our series. The real output gap was then calculated by the difference of real GDP from the real GDP-trend represented by the HP trend.

With respect to fiscal policy we want to distinguish between expansionary and restrictive fiscal policies. We decided to focus on the general government including social securities as revenue and expenditure developments in the social security system are subject to political discretion and tend to affect the overall fiscal stance of the government, which in turn influences macroeconomic development. However, we have excluded interest spending because the government's ability to change this spending category is very limited. This is in line with other approaches to fiscal policy analysis (see e.g. Perotti (2004)), but not uncontested as for example Blanchard and Perotti (2002) include interest spending. It should be noted here, that the general timing pattern of fiscal policy over the business cycle is only slightly affected by interest

spending, which indicates that we could expect very similar results even if we integrate interest spending. For our series of the real fiscal stance of the government we have subtracted general government expenditures (excluding interest) according to the national accounts definition from the general government revenues (mostly taxes and social security contributions). To derive a measure of the real and seasonal adjusted fiscal stance we have first realized the expenditure and the revenue series with the chain-linked GDP deflator and then adjusted the series seasonally.

Figure 1: Fiscal Policy over the business cycle in Germany 1970-2008

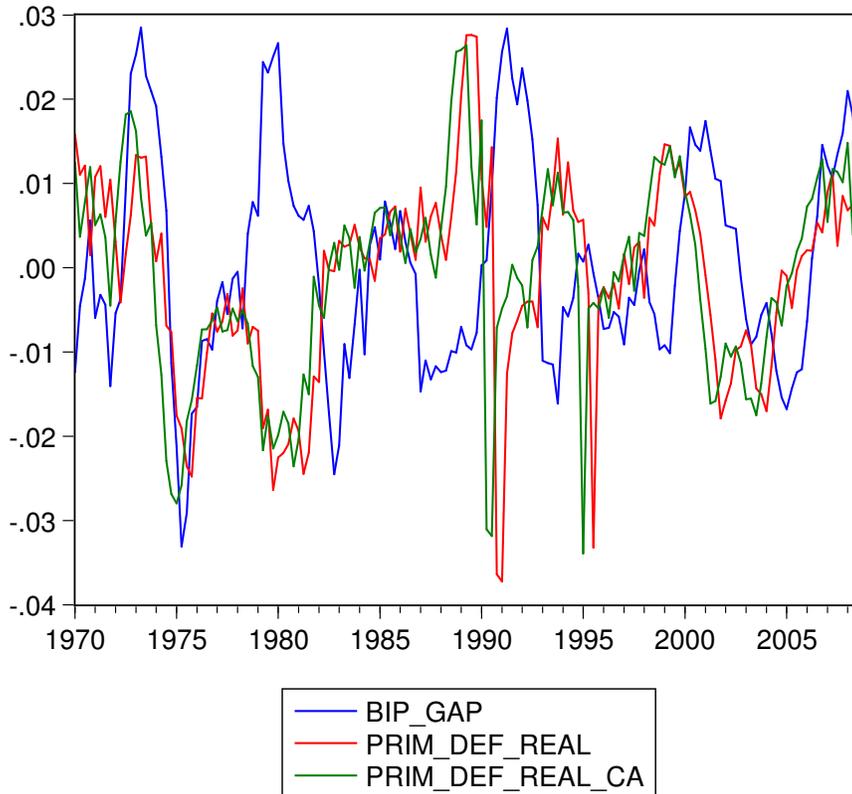


Figure 1 shows the development of the business cycle (real GDP gap - blue line) and the government fiscal stance as the primary deficit (red line). We see already that the two series seem to move very closely together in some periods, while the series diverge strongly during other periods of time. This can be seen as a first indication that a time-invariant analysis might not be the optimal choice in this case.

Discretionary fiscal policy is usually studied based on cyclically adjusted fiscal data, which results from a filtering of the automatic stabilization processes from the data. A "cyclical" component is calculated and then subtracted from the aggregate

revenue and expenditure developments. This cyclical component is calculated based on an indicator for the state of the business cycle (usually the output gap or the gap of one of its components) and an elasticity measure for the effects of business cycle fluctuations on fiscal developments. We build our analysis here on the standard OECD method of cyclical adjustment, laid out e.g. in Girouard/André (2005). The elasticity of the general government deficit for Germany is calculated to equal 0.49 in 1996, 0.47 in 2000 and 0.51 in 2003 (see Girouard/André (2005), p. 24), meaning that an output gap of -2 percentage points would account for a cyclical general government budget deficit of 1 percentage point. However, these calculations are all for annual data, while we use quarterly data here, which makes assumptions about the quarterly lag-structure of the elasticities reflecting the automatic stabilizers necessary. We derived these quarterly elasticities by the following steps. First, we assumed that the aggregated yearly elasticity of the general government budget deficit with respect to the real GDP gap equalled during the period of our analysis 0.49 - the average of the annual values calculated by the OECD for 1996, 2000 and 2003. In a second step we analysed the correlation of the GDP gap and the fiscal stance of the government in the same period, which turned out to be on average 0.155 over the whole sample. In line with other studies on the reaction of fiscal policy to business cycle developments (see e.g. Blanchard/Perotti (2002)) we argue that discretionary fiscal policy cannot react in the very same quarter to changes in macroeconomic development. Therefore we assign the direct contemporary reaction of the fiscal stance to changes in the GDP gap (which equals on average 0.155 throughout our sample) completely to the automatic stabilizers. Automatic stabilizers work strongest with respect to tax revenues (see e.g. Girouard/André (2005)) and here the effect can be expected to be strongest in the quarter after the change of the macroeconomic conditions as tax payments of the cyclically sensitive profit taxes are usually adjusted on a quarterly basis. Therefore we argue that the effect of automatic stabilizers should be slightly higher in the first quarter than contemporary and fade out afterwards. Based on these arguments and additional applications of the methods outlined in Girouard/André on the quarterly data we applied elasticities to the real GDP gap to calculate the effect of the business cycle on the fiscal stance which equal 0.155 (contemporary reaction, in period t), 0.17 (in $t+1$, the quarter after the change of the macroeconomic conditions) and 0.1 (in $t+2$). Taken together these three elasticities add up to an aggregate elasticity of 0.49 on an annual basis. The yellow line in figure 1 reflects the fiscal stance which was cyclically adjusted applying these elasticities and the output gap calculated as outlined above. We see that cyclical adjustment does affect the fiscal stance quantitatively and qualitatively (sometimes a deficit becomes cyclically adjusted a surplus) but does not change the complete pattern observable in the time series. In the empirical part we

will usually estimate the models with cyclically adjusted and cyclically unadjusted data, which gives us the possibility to discuss the effects of cyclical adjustment on our results.

4 Evidence based on time-invariant VAR analysis

Our aim is to analyze the timing of fiscal policy over the business cycle - an attempt, which is directly affected by the possible endogeneity of fiscal policy and the business cycle: not only the business cycle might affect fiscal policy but fiscal policy might affect the business cycle as well. To tackle the general endogeneity problem and to account for the fact that fiscal policy develops over time we propose modern Vector-Autoregressive (VAR) time-series analyses in which such relationships can be modelled by vector autoregressions in an intuitive way. This is as well an important reason why VAR analyses have become the dominant method to study the interactions fiscal policy and the business cycle.

Here we start our analysis with a time-invariant VAR model as a background for our time-variant estimations in the following parts.

4.1 The benchmark model: The reduced form time-invariant parameter VAR

The benchmark reduced-form VAR(p) with time-invariant parameters is given by

$$y_t = \nu + A_1 y_{t-1} + \dots + A_p y_{t-p} + e_t, \quad t = 1, \dots, T,$$

where the residuals e_t follow a zero mean noise process with a constant residual variance covariance matrix $E(e_t e_t') = \Sigma_e$. Moreover, we assume independence for e_t and e_s , if $s \neq t$. The sample size is T time series observations and p presample values for each variable. The vectors y_t , ν and the residual vector e_t are of dimension $n \times 1$, whereas the lag coefficient matrices are all $n \times n$. For one observation t this model can be rewritten in a more simple form,

$$y_t = B Z_{t-1} + e_t, \quad t = 1, \dots, T$$

where the $n \times 1 + np$ -coefficient matrix is defined by $B = [\nu : A_1 : \dots : A_p]$, and the $1 + pn \times 1$ regressor matrix is given by $Z_{t-1} = \begin{bmatrix} 1 & y_{t-1} & \dots & y_{t-p} \end{bmatrix}$. Vectorizing this model and applying some Kronecker rules then leads to the following form of the VAR(p) process to which we will refer from now onwards:

$$y_t = Z_t \beta + e_t, \quad t = 1, \dots, T$$

where the $n \times m$ matrix Z_t includes all the information from the regressors, $Z_t = \left(Z'_{t-1} \otimes I_n \right)$,

and the time invariant $m \times 1$ - vector β the coefficient values for the n constants and pn^2 lag coefficients $\beta = \text{vec}(B)$. In this version of the n -dimensional VAR of order p the regressor matrix is not restricted and contains only a constant as exogenous explanatory variable.

4.2 Identification of an output gap shock

To analyse the structural relations between the output gap and the fiscal stance of the government based on reduced form VAR processes we assess the impact of exogenous and unanticipated shocks as they do not affect the systematic and endogenous economic relations between the aggregates.¹ There are two essential problems of such a strategy. **First**, one has to make sure that reactions in the variables displayed e.g. by impulse response functions are in fact due to the considered economic shock. Therefore the shock must work isolated from other potential shocks. This is usually the case if the disturbances of the estimated system are instantaneously uncorrelated and thus the residual variance-covariance matrix of the estimated process is diagonal. Isolation can for example be achieved by a transformation of the reduced form disturbances so that the residuals become uncorrelated and imply a specified contemporaneous structure of the economy. This leads to a structural parameter model that can be written in the following general AB-form (without any additional deterministic terms),

$$A y_t = A A_1 y_{t-1} + \dots + A A_p y_{t-p} + B e_t, \quad t = 1, \dots, T$$

with $A^{-1} B e_t = u_t$,

and hence $\Sigma_u = E(u_t u_t') = A^{-1} B E(e_t e_t') B' A^{-1'} = A^{-1} B \Sigma_e B' A^{-1'}$,

where $(A^{-1} B)^{-1}$ is a $n \times n$ variance-covariance decomposition matrix, such that Σ_u is a n -dimensional diagonal matrix with the variances of the isolated or respectively orthogonalized errors that now indicate economic innovations (or respectively structural shocks) on the main diagonal.

But even if a parameterization of the structural AB model is chosen in such a way that the error terms are not instantaneously correlated, the impulse responses based on the reduced form VAR may still not be unambiguously determined (our **second**

¹Structural shocks and also the identification scheme are assumed to be time-invariant over the whole sample. Only the lag and constant term coefficients are allowed to be time varying later on.

problem). In general, there exist more than one parameterization of the matrices A and B which lead to a diagonal variance-covariance matrix of the structural shocks.² In other words, more than one economic structure can yield a diagonal variance-covariance matrix of the innovations. Therefore we have to find a parameterization of the structural model that leads to isolated shocks and at the same time represents a reasonable economic structure.

To keep the schemes simple we apply the commonly used assumption that identification of the system can be achieved by implementing an economic structure only on the contemporaneous interactions of the reduced form residuals and not on the contemporaneous relations of the variables themselves, which corresponds to setting $A = I_n$ (Bernanke, 1986). This leads to the B-SVAR form to which we will refer from now onwards for both datasets,³

$$Be_t = u_t.$$

In a next step we use standard orthogonalization procedures in combination with institutional information about the timing of the German tax and transfer system and the existence of political decision lags to identify an isolated structural shock in the output gap. More precisely, we used the Cholesky-type decomposition approach introduced by Sims (1980) to orthogonalize the shocks and check whether this leads to a reasonable economic structure which seems to be the case.⁴ The implied parameterization leads to a recursive structure of the economy which is in line with the usually applied time lags in the German tax and transfer system and evidence for political decision lags.⁵ Based on the variable ordering in our VARs the fiscal stance reacts contemporaneously to the output shock u_t^{gdp} , but there is no feedback reaction from the structural primary deficit to the output gap within the same quarter.⁶ This

²There may exist more than one solution to the linear system of equations.

³In general, empirical studies that consider fiscal policy issues refer to one or more of four main identification approaches (Caldara and Kamps, 2008). First, the standard recursive approach introduced by Sims (1980) and applied in the context of fiscal policy analysis by Fatas and Mihov (2001); second, the structural VAR approach by Blanchard and Perotti (2002) and Perotti (2005, 2007); third, the sign-restriction method by Uhlig (2005); and, fourth, the event-study method by Ramey and Shapiro (1998). Alternatively, there exist approaches that distinguish between short- and long run structural shocks (see e.g. Blanchard and Quah, 1989, or Lee and Chin, 2006). However, in this study we will only focus on the effects that structural shocks have in the short run in a simple B-SVAR. In future research we plan to apply the approach by Blanchard and Perotti in an AB-SVAR framework using cyclically unadjusted data as well.

⁴In fact, this approach is most frequently used in fiscal and monetary policy applications of VARs and TVP-VARs, see e.g. Muscatelli et al., 2002, Muscatelli et al. (2007) or Baumeister, Durinck and Peersman (2008).

⁵In fact the recursive structure of the economy is given by the recursive relations of the structural residuals, u_t .

⁶Applying the Cholesky-type decomposition of the residual variance-covariance matrix, the matrix B^{-1} is now lower-triangular with unit diagonal and called the Cholesky-decomposition matrix. The orthogonalized system is just-identified as only one additional zero restriction is set.

is in line with the literature that studies the interaction of fiscal policy and the business cycle (see e.g. Blanchard and Perotti (2002) or for Germany e.g. Heppke-Falk et al. (2006) and as well the discussion in part (III)). In our approach the contemporaneous reactions of the fiscal stance to changes in the output gap can be assigned completely to the working of the automatic stabilizers - as reactions in discretionary policy do not occur only with a time lag.⁷ We compute the contemporaneous reactions by a maximum likelihood estimation using the scoring algorithm by Amisano and Giannini (1997) for a general B - model with constant parameters, where we set a no-contemporaneous-feedback zero restriction. This method leads to an reaction of 0.155(0.101) of the fiscal stance within the first quarter to a one percent output shock.⁸ The Cholesky decomposition delivers exactly the same value as it implies the same zero restriction.

Hence, under the assumption that there is no instantaneous reaction in discretionary policy, the contemporaneous reaction of the cyclically adjusted fiscal stance to a Cholesky identified GDP gap shock should be equal to zero.⁹ If not, the applied cyclical adjustment procedures might be false. Nonetheless, the question how automatic stabilization works within the following quarters is more complicated (see as well the discussion in part III).

To analyse the economic impact of a shock in the business cycle over a certain horizon we will use impulse responses, that can be interpreted as forward-looking policy reaction functions.

4.3 Benchmark results

We used multivariate least squares (LS) estimation to obtain time invariant values for the coefficients of the VAR(p) based on cyclically unadjusted and cyclically adjusted data. The optimal lag order was chosen to be according to the Schwartz (SC) information criterion, autocorrelation analysis and due to the fact, that the

⁷However, the assumption that there is no contemporaneous reaction in discretionary fiscal policy stands in contrast with the assumption of Heppke-Falk, Tenhofen and Wolf (2006) and Blanchard and Perotti (2002). In their approach any contemporaneous and ongoing reaction in a fiscal aggregates represent only changes in a discretionary policy of the government. But there is another strong restriction in their approaches. To identify a discrete fiscal policy shock, they have to assume that automatic stabilization only takes place within the first quarter. They estimated the contemporary reactions of the fiscal aggregates (in fact only revenues) and the n explicitly correct for this reaction in the revenue component by setting the corresponding substract-restriction into the A matrix of their AB-SVAR. However, in our opinion this assumption is rather stronger than our assumption of a policy decision lag in a parliamentary and federal political system.

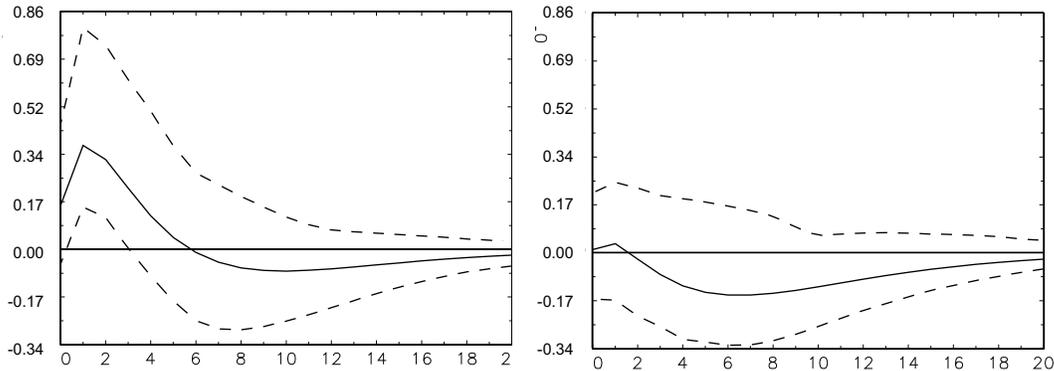
⁸Alternatively, using a general AB model with diagonal B matrix and a restriction that ensures no contemporaneous feedback from the fiscal stance to the output gap, the contemporaneous elasticity between the variables themselves can be estimated by ML using the scoring algorithm. Here we estimated a value of 0.159 (0.0976), which differs only slightly from the alternatively derived results.

⁹This can also be tested using the ML estimation procedure or simple tests on instantaneous causality. In such a case a forecast error identification scheme can be applied.

time-invariant VARs will be taken as reference for the time varying parameter models that may be over-fitted with higher lag-orders.

Figure 2 shows the impulse-response functions for cyclically unadjusted (left panel) and cyclically adjusted data (right panel) with 95% confidence intervals. For the unadjusted data we see a positive contemporary reaction of 0.16 which reflects the working of the automatic stabilizer. The effect of the business cycle on fiscal policy becomes significant in the first quarter and increases to 0.35. It continues to be significant until the second quarter but fades out afterwards.

Figure 2: Impulse responses benchmark model



With respect to the cyclically adjusted data the contemporary effect is zero. From the second quarter on the effect is negative and it is insignificant for the whole period analyzed. The results of the forecast error identification approach for cyclically adjusted data look similarly and are presented in appendix A iii.

If these specifications would be adequate the results would indicate that discretionary fiscal policy (not including the automatic stabilizers) tended to be a-cyclical in Germany from 1970 to 2008 - a result largely in line with the findings of other analyses of cyclically adjusted data in European economies (see for a discussion part 2). However, the specification and the results including the confidence intervals are only reliable in case of parameter stability. If the response pattern of fiscal policy to the business cycle has changed over time, this would be reflected in indications for parameter instability. Recursive coefficient estimates by multi-variate least square estimation (results presented in appendix A i) and ii)) speak strongly in favour of parameter instability. Therefore we proceed with a VAR analysis, which is potentially more appropriate as it allows for a time-varying of the parameters

5 A time-varying parameter (TVP) VAR approach

The evidence for parameter instability in time-invariant models (see appendix A) speaks in favour of a time variant model, where VAR coefficients and the corresponding impulse responses might differ over time. One way to analyse these changes would e.g. be to simply build sub-samples for each fiscal policy or business cycle regime. But this would also lead to very short samples of only a few observations, which reduces the reliability of the estimates significantly. Alternatively, there is a large literature in macroeconomics which documents structural breaks and other sorts of parameter changes in many time series variables (see e.g. Stock and Watson, 1996). Moreover, a wide range of alternative specifications have been suggested, including Markov-switching VARs (e.g. Paap and van Dijk (2003), or Sims and Zha (2006)) and other regime-switching VARs (e.g. Koop and Potter (2006)). However, the most frequently used method is to use time varying parameter VARs (see e.g. Doan, Sims and Litterman (1984), Primiceri (2005) or Muscatelli, Spinelli and Trecroci, (2007)).

5.1 The reduced form (homoskedastic) time-varying parameter VAR

In this part of the study we will consider a variation of the benchmark VAR model that was used to explain the data generation process in the last section. According to the results from parameter stability analysis in section 2 there might be better models to explain the data and the structural relations between the variables. But how could these kind of models look like?

Taking into account the *changing regimes hypothesis* we have identified in the beginning and further the information about switching parameters we want to allow structural changes in policy regimes to evolve gradually over time. An adequate empirical solution to capture these properties would be a VAR model where the coefficients can potentially change over time.¹⁰ We will focus on an “time varying parameter” (TVP) VAR in the following parts of the study.¹¹ We introduce a VAR framework, where

¹⁰There are several other classes of adequate of models and possibilities that take into account the consequences from the intuitive appeal and switching parameters. Building sub-samples for each fiscal policy or business cycle regime is one of them. But this would also lead to very short samples of only a few observations which reduces the reliability of the estimates significantly. Alternatively, there is a large literature in macroeconomics which documents structural breaks and other sorts of parameter change in many time series variables (see e.g. Stock and Watson, 1996). Moreover, a wide range of alternative specifications have been suggested, including Markov-switching VARs (e.g. Paap and van Dijk (2003), or Sims and Zha (2006) and other regime-switching VARs (e.g. Koop and Potter (2006)). However, the most frequently used method is to use time varying parameter VARs (see e.g. Doan, Sims and Litterman, 1984, Primiceri, 2005, or Muscatelli, Spinelli and Trecroci, 2007).

¹¹However, we will not consider concerns that might result from time varying moments of the processes, covered in e.g. in cointegration approaches based on trended series. Instead, we will continue working with the detrended and seasonally adjusted series from the first parts.

the VAR parameters evolve over time as observations are added. It should be noted, however, that only the parameters but not the structure of error covariance matrix can potentially change over time (so-called homoskedastic TVP-VARs). Some **general properties** of our n -dimensional homoskedastic TVP-VAR(p) process including an intercept term are presented below.

Starting from the vectorised form of the time invariant VAR(p) model given in part 4.1, we will now introduce a version with time-dependent coefficients,

$$y_t = Z_t \beta_t + \varepsilon_t, \quad t = 1, \dots, T$$

where the residuals ε_t follow a zero mean noise process with **time-invariant** residual variance covariance matrix $E(\varepsilon_t \varepsilon_t') = H_t = H$. We assume independence for ε_t and ε_s , if $s \neq t$. The sample size consists of T time series observations and p presample values for each variable. The vectors y_t , ν_t and the residual vector ε_t are of dimension $n \times 1$, whereas the lag coefficient matrices are all $n \times n$. The regressor matrix is of dimension $n \times m$ and is given by

$$Z_t = (Z'_{t-1} \otimes I_n).$$

Now, the $m \times 1$ - vector β_t contains the values for the n constants and pn^2 lag coefficients in period t ,

$$\beta_t = \text{vec}(B_t).$$

Again the regressor matrix of the n -dimensional time varying VAR of order p is not restricted and contains only a time varying constant as exogenous explanatory variable matrix.¹²

5.2 The normal linear state space representation of the TVP-VAR

One major problem we - as every empirical application - face is that relations between the economic variables (reduced and structural forms) can not be observed or measured directly. In our case we do only know that the relations can possibly change over time and that the data can be analyzed by a TVP-VAR. In such a situation a so called state space representation of our model with time varying parameters can make life easier as it includes some additional information on the stochastic behaviour of the reduced form coefficients over time.

¹²we plan to integrate other exogenous variables such as dummies and fiscal variables identified as exogenous in the next versions of the paper.

The main idea of state space models is that the VAR coefficients – from now on called states- can be calculated recursively from measurable data described by a certain data generating process (e.g. by a TVP-VAR), while assuming that they follow a stochastic process with known properties.¹³ Hence, usually a state space model is a two-layer model, where the external layer involves the measurable data and is therefore called measurement equation and the internal layer involves the information of the motion of the states and is therefore called state equation. More formally and according to Ltkepohl (2006, p. 611), this general idea can also be expressed as the dependency of an observable and possibly multiple time series y_1, \dots, y_T upon an unobservable state z_t that is driven by a stochastic process, whereas the dependency between y_t and z_t is described in the measurement equation,

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t, \quad t = 0, \dots, T.$$

In our case this equation takes the form of the TVP-VAR from above. In other words the measurement equation describes our data generating process (DGP), where y_t is the $n \times 1$ vector of observations, α_t is the unobservable $m \times 1$ state vector at time t and ε_t the $n \times 1$ vector of serially uncorrelated zero mean measurement errors that have a time invariant variance-covariance matrix, H . H , the $n \times m$ regressor matrix Z_t , and the $n \times 1$ vector d_t (that possibly contains other time invariant states and their regressors are called the measurement system matrices).¹⁴

Moreover, we call the unobserved stochastic process of regression coefficients, α_t , the state (or transition) equation. This process is assumed to evolve like a multidimensional (m-dimensional) AR(??)-process of the form

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t, \quad t = 1, \dots, T$$

where the $m \times m$ matrix T_t is called the transition matrix that involves all the information of how past states enter the measurement equation at time t . The $m \times 1$ vector c_t involves other exogenous components such as a constant or dummy variables and R_t is a matrix of dimension $m \times g$ that involves structural relations between the disturbances of the states that are described in the $g \times 1$ vector η_t and that have a time invariant $m \times m$ variance-covariance matrix, Q . Q , T_t , c_t and R_t are called the state system matrices.¹⁵

¹³State space models are frequently used in other sciences. Good examples can be found especially in the literature of physical sciences and engineering.

¹⁴The $n \times 1$ vector d_t can be decomposed into the $n \times r$ matrix W_t that could contain lagged dependent or other explanatory variables with time invariant coefficients, and the corresponding $r \times 1$ vector δ_t with the time invariant parameters.

¹⁵In the economic literature on state space models the movement of the state vector is generally assumed to be a first-order autoregressive process (see e.g. Hamilton (1994), Doan, Litterman and Sims (1984) or Harvey (1992)).

To keep the structure of our state space model as simple as possible we make **four additional assumptions** for our specific version of the state space model (see e.g. Tanizaki, 1996):

1. We have already assumed only linearized relations in the measurement and state equation.
2. We further assume the measurement errors and the disturbances of the state equation each to be serially uncorrelated and to follow a multivariate normal distribution with mean zero and time invariant variance-covariance matrices H and Q .¹⁶ Furthermore, they are uncorrelated with each other and some initially chosen normally distributed state-variables, α_0 , i.e. $E(\varepsilon_t \varepsilon'_s) = 0$, $E(\eta_t \eta'_s) = 0$, $E(\varepsilon_t \eta'_s) = 0$ for all t and all s , and $E(\varepsilon_t \alpha'_0) = 0$ and $E(\eta_t \alpha'_0) = 0$ for $t = 1, \dots, T$. The last assumption excludes correlation between ε_t and α_t as well as between η_t and α_{t-1} ($E(\varepsilon_t \alpha'_t) = 0$ and $E(\eta_t \alpha'_{t-1}) = 0$ for all t).¹⁷
3. Besides the time invariant system matrices H and Q , we further assume the other system matrices to be time invariant and the vector d_t to be equal to zero, $Z_t = Z$, $d_t = d = 0$, $T_t = T$, $c_t = c$ and $R_t = R$, whereas $R = I_m$ as g equals m . In order to be able to estimate all the remaining system matrices based on e.g. procedures that evaluate the likelihood of corresponding models all the remaining system matrices Z , T , H and Q , have to be assumed to depend on an unknown vector of parameters that includes randomly drawn coefficient values, which determine the system matrices. To keep our application as simple as possible we only consider the residual variance-covariance matrices of both (the measurements, H , and the states, Q) to depend on a set of randomly drawn and normally distributed parameters all contained in the vector PAR . The residual variances-covariances might be of most interest concerning the evaluation and structural analysis of our estimated state space model. Moreover, alternative and even less comprehensive scenarios in which only one of the variance-covariance matrices, H or Q , or even none of them depends on the randomly drawn set of parameters, PAR , are considered. The smaller the set of parameters, the less the computational burden of the application.
4. In addition, we make some slightly more specific and economic assumptions on the characteristics of the linear movement process of the states. According to

¹⁶The independence assumption of the disturbances in the state equation is equivalent to setting g equal to m .

¹⁷According to Tanizaki (1996) the normality assumption of the disturbances is required, if we want use the standard linear recursive algorithm of the Kalman Filter that is derived based on density functions, to estimate the states. This is discussed more detailed in the next section.

Doan et al. (1984) and Koop (2009) we assume the state vector α_t to be driven by a specific version of multivariate autoregressive process of order one that can easily be transformed into standard processes such as random walks. Thereby we assume the state vector in each period t to depend on a weighted sum of its initially defined steady state values. Thus, the state in period t , α_t is modelled as an initially defined steady state value a_0 plus the weighted deviation of the last periods state value, α_{t-1} , from its steady state plus a zero mean residual component,

$$\alpha_t = a_0 + T(\alpha_{t-1} - a_0) + \eta_t, \quad t = 1, \dots, T,$$

which corresponds to our above notation of the model with a constant term that equals the weighted steady state of the state vector, $c = (1 - T)a_0$.¹⁸

From an economic perspective this representation is interesting as the transition matrix T can now be referred to as a **“rate of decay towards the prior mean”** or respectively towards the **prior steady state** (Doan et al. (1984), p. 8).¹⁹ The higher the rate of decay, the weaker is the impact of the initially set steady state on the value of the state vector in each period t and thus, the more dynamic reactions in the states are possible.

In order to check our results for robustness we apply different values for the rate of decay. Depending on the choice of T , the randomly varying coefficients may follow different stochastic processes.

According to Koop (2009) it is common in economic applications to restrict the states to follow a stationary process and therefore impose the restriction that $|T| < I_m$. This stationarity assumption for the states seems to be reasonable for our application As well. Based on the narrative and descriptive analysis of the data in section 2 we expect the fiscal regimes to possibly change randomly and frequently over the long horizon from 1970 to 2008, instead following any kind of a long run trend.²⁰ Obviously,

¹⁸In fact, this state equation can be rewritten as $\alpha_t = (1 - T)a_0 + T\alpha_{t-1} + \eta_t$ or in terms of our notation where the constant term is represented by c as $\alpha_t = c + T\alpha_{t-1} + \eta_t$, where $c = (1 - T)a_0$.

¹⁹The assumption of a steady state fiscal stance level seems to be reasonable with respect to recent studies from dynamic stochastic general equilibrium theory. This steady state level can be used as the mean of the prior distribution in the prior setting procedure. In such a case a state equation of the underlying form would ensure that deviations from the steady state values in the state coefficients are expected to decay, depending on how high the rate of decay is chosen. In expected values, a high rate of decay in combination with a large deviation of the last periods states from the steady state level would lead to a current state vector far away from the steady state assuming the shock to be zero in expectations. Thus, a high rate of decay leads to persistence in the deviations in each direction. A lower rate leads to lower and less fluctuations around the steady state. In general the rate of decay is assumed to be $0 < |T| \leq I_m$. We experimented with values between 0.1 and 1.

²⁰In a reference scenario using LS priors where only the covariance matrix of the measurement residuals has to be estimated by ML, a rate of decay below $T = 0.98$ ceteris paribus does lead to very smooth state coefficients in our study. Thus, the impact of the initially set steady state (prior mean)

this stands in contrast with any specification of non-stationary states as they can be modelled by e.g. random walks with drift.²¹

However, also $T = I_m$ can imply stationary states and is frequently used in the economic TVP-VAR literature (see e.g. a number of experiments in Doan et al. (1984) or Primiceri (2005) for a case where the volatility of a process is time-varying and is driven by a stochastic process which is a variation of a random walk). In fact, setting $T = I_m$ implies that α_t differs from α_{t-1} only by an additive zero mean random component. In such a case the states are driven by a multivariate random walk over time that also seems to be a reasonable specification as frequent and random regime changes can be modelled. The main difference between this version and the version where $|T| < I_m$ is that there is at least no impact of the initially set steady state and thus frequent dynamics in the states are possible.²²

Based on these main assumptions we will refer to the following TVP-VAR(p) in state space representation from now onwards:

$$\begin{aligned} & \text{(measurement equation)} \quad y_t = Z\beta_t + \varepsilon_t, & t = 0, \dots, T \\ & \text{(state equation)} \quad \beta_t = b_0 + T(\beta_{t-1} - b_0) + \eta_t, & t = 1, \dots, T. \end{aligned}$$

where $\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} H & 0 \\ 0 & Q \end{pmatrix} \right]; \alpha_t = \beta_t; a_0 = b_0$

is very strong even under conditions where the rate of decay is very high. Usually this can also result from very tight prior variances-covariances for the state variables that are used in the initialization. Therefore we tried values between 10^{-6} up to 10^6 times the OLS variances as the variance of the prior distribution, but this does not really decrease the smoothness of the of the states around their steady state. When more volatile states occurred, they were usually combined with dramatically higher standard deviations. In some cases playing around with the tightness also led to incredible huge or small impacts of shocks in the output gap on the fiscal stance. Another possibility to reduce the smoothness in the states might achieved by setting less “tightness on the state residuals”. In other words the values of the main diagonal elements of the state residual variance covariance matrix are increased (later on we will see that this corresponds to higher values for λ). However, implementing values significantly higher than the ones considered in the literature by e.g. Doan, Sims and Litterman (1984) or Muscatelli et al. (2007) often leads to exploding or incredible large impulse responses in many periods. Therefore we will stick to the tightness on the state residuals that is established in the literature. In the $|T| < I_m$ scenario we refer to a rate of decay equal to $T = 0.999$, to be able capture the dynamics in the states, as suggested by Muscatelli et al. (2007) for a similar model specification with exogenously determined residual covariance matrices. By the way, Muscatelli et al. (2007) do not find differences for rates between 0.3 and 0.999 (they only show the results for 0.999).

²¹In a drifted model frequent breaks could only be introducing by additional structural break components into the trend slope and the constant. A random walk with drift could be introduced by e.g. setting the transition matrix equal to I_m and additionally introducing another constant term ($c_2 \neq 0$)

²²However, no significant differences can be found between both reasonable scenarios suggested in the literature, $|T| < I_m$ with $T = 0.999$ and $T = I_m$ for the reference scenario based on LS priors. Therefore we will focus just one of the versions from now on. The version with $T = 0.999$ assuming stationary states and the state vector in each period t to depend on a weighted sum of its initially defined steady state values and a random component.

5.3 Identification of an output gap shock in the TVP-VAR

The identification scheme from the constant parameter case is transferred to the case of a TVP B-SVAR model. Thus,

$$B\varepsilon_t = u_t,$$

where $G = E(u_t u_t') = BE(\varepsilon_t \varepsilon_t')B' = BHB'$,

and H is the ML estimated variance-covariance matrix of the measurement residuals and B^{-1} is a $n \times n$ lower triangular Cholesky-type decomposition matrix with unit diagonal, such that G is a n -dimensional diagonal matrix with the variances of isolated or respectively orthogonalized errors that now indicate economic innovations (or respectively structural shocks) on the main diagonal.²³

5.4 Estimation strategy: Maximum likelihood estimation of the state space model - A three stage estimation procedure

Our estimation strategy consists of three main steps and is based on conventional methods of developing models for econometric time series. First, we use Kalman filter recursions given some initially defined variance-covariances of the measurement and state disturbances and the defined priors to estimate the state variables for each point of time up to $t = T$. Based on this, we evaluate the log-likelihood function starting with the initial set of (estimated) hyperparameters to find the residual variance-covariance estimators that maximize the likelihood of the corresponding models for changing sets of hyperparameters. Given the optimal variance-covariance estimators and the initially defined priors we finally compute the optimal Kalman filtering estimates for the states and their variance-covariance matrices.

Obviously, the core of the applied 3-Step ML estimation procedure is the evaluation of the likelihood function. Thereby, the log-likelihood function is evaluated conditional on the available data $y = (y_1, \dots, y_T)'$ with an iterative algorithm to find the optimal estimators for the residual variances-covariances of the measurement equation, H^* , and - depending on the scenario- also of the state equation, Q^* , given some residual variance-covariances, H and Q , the corresponding filtering estimates $b_t = \beta_{t|t}$ and $P_t = P_{t|t}$ for $t = 1, \dots, T$, for each randomly taken normally distributed set of parameters, PAR , and some prior settings for the states, $\beta_0 \sim N(b_0, P_0)$.²⁴

²³As the variance-covariance matrix of the measurement disturbances is time-invariant in our application we do not need to implement a time-varying Cholesky decomposition.

²⁴In this paper we focus on versions where only the variance covariance matrix of the measurement errors has to be estimated as only this variance covariance matrix is set to depend on some randomly drawn and normally distributed parameters, PAR . The variance-covariance matrix of the state

The **log-likelihood function** of our Gaussian model can be derived in three steps.²⁵ In a first step, using Bayes' theorem and the sample density function, the joint density function can be derived, where $\theta = (H, Q, b_0, P_0)$ is the vector of hyperparameters and $f = f(y_t | Y_{t-1}; \theta)$ the distribution of y_t , conditional on the information set at time $t - 1$, Y_{t-1} . Additionally taking the Gaussian properties of our model into account the true state vector of our state space model at time t is normally distributed by assumption with mean b_t and variance-covariance matrix P_t . Therefore, also y_t is normally distributed with recursively given mean $E(y_t | Y_{t-1}) = y_{t|t-1}$ and variance-covariance matrix $Cov(y_t | Y_{t-1}) = F_t$. In a second step the Kalman filtering equations can be used to estimate these quantities, given a specific vector of hyperparameters $\theta = (H, Q, b_0, P_0)$. Thus, $y_{t|t-1} = Zb_{t|t-1} + d$ and $F_t = ZP_{t|t-1}Z' + H$.²⁶ Finally, taking the joint density function, the normality assumption and the information from the Kalman filtering recursions, the log-likelihood function of our Gaussian state space model is the following:

$$\ln L(\theta | y) = \ln L(H, Q, b_0, P_0 | y) = -\frac{nT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |F_t| - \frac{1}{2} \sum_{t=1}^T \nu_t' F_t^{-1} \nu_t,$$

where we denote the dimension of our data matrix by a general dimension n as we consider different dimensions of the TVP-VAR and ν_t is the estimation error from the Kalman filtering procedure. This function has to be maximized with respect to the vector of hyperparameters $\theta = (H, Q, b_0, P_0)$, where the factors b_0 and P_0 are the initially set priors and thus, only H , and - depending on the scenario - as well Q have to be estimated.²⁷ This is an unconstrained nonlinear optimization problem that can be solved using a standard iterative algorithm.²⁸ In fact, the procedure

residuals is exogenously specified. Nonetheless we also estimated versions in which we assumed the variance-covariances of the residuals in the state and the measurement equation H and Q , both to be dependent on some randomly chosen and normally distributed parameters. In these more comprehensive versions the optimal residual variance-covariance matrices are both estimated by evaluating the likelihood function of the model conditional on the data and given some initially estimated state vectors as initial set of hyperparameters to run the procedure. However, the main problem of such larger versions is that due to the large number of parameters to estimate the system is always in danger to become overparameterized. With respect to our relatively small sample of around 154 (observations depending on the specification) overparameterization is a serious risk. Indeed the ML estimations lead to rather incredible state coefficients and in many cases to exploding impulse responses at some points of time. In short, the versions where only the measurement covariance matrix is estimated lead to significantly more stable results.

²⁵A more detailed derivation of the joint density function is shown in Appendix B. The derivation is based on Ltkepohl (2006).

²⁶In our case the vector d is equal to zero.

²⁷Some initial residual variance-covariance matrices for the measurement and state disturbances are constructed by from some randomly drawn normally distributed parameters in order to be able to start the procedure.

²⁸We decided to use a direct search method called Nelder-Mead simplex search method that is

we implemented is based on the dual approach of the optimization problem, the minimization of the negative log-likelihood function.

As mentioned the **Kalman filter recursions** are a useful tool in the evaluation of the likelihood function because they provide an optimal estimator b_t for the states β_t at each period of time, given under normality assumptions and the measurements y_1, \dots, y_T . The filtering estimates resulting from the updating (or correction) step at each point of time t are given by the expected value of the state variable conditional on the observations up to this point of time, $b_t = \beta_{t|t} = E(\beta_t | y_1, \dots, y_t)$ for $t = 1, \dots, T$. Once our iterative algorithm converges to an optimum, H^* and Q^* , the Kalman filter recursions are again used to compute some optimal estimates for the states, b_t^* in a final step.²⁹ However, to initially start the Kalman filter recursion procedure some initial states have to be defined. In other words, a **prior distribution** described by mean and covariance for the initial states has to be chosen exogenously at first.

5.5 Initialization: Priors

In the last section we have introduced the Kalman filter as a useful and efficient tool to recursively compute an optimal estimator b_t for the states β_t at each period of time, given the measurements y_1, \dots, y_T and under normality assumptions. To start the Kalman filter recursion procedure a prior distribution for the states described by mean and variance-covariance matrix has to be chosen exogenously at first. As the initialization potentially has some influence on the estimation procedure and the posterior states while it leaves some degrees of freedom to the analyst, we decided to use different initialization scenarios.

documented in Lagarias, Reeds, Wright, and Wright (1998). This method does not use numerical or analytic gradients in order to keep the solution procedure simple and robust. In general, the chosen procedure can handle discontinuity, particularly if it does not occur near the solution. Alternatively, more efficient numerical methods as gradient or respectively scoring algorithms may be used. However, a scoring algorithm might have poorer convergence properties far from the optimum and has a high computational burden. Apart from the choice of an optimization algorithm, every procedure faces the problem that it might only converge to a local minimum.

²⁹In the Kalman filtering recursions we used the slightly modified method by Anderson and Moore (1979) to compute the variance-covariances of the filtered states in order to avoid negative definite matrices because of round-off errors. One essential problem of the Kalman filter is that for every estimate only the information up to time period t is used. In some cases more reliable estimates may be reached when all the sample information up to time T is taken into account at each iteration. Thus, the so called Kalman smoothing recursions may produce more reliable results than the filtering recursions. Nonetheless, also the smoothing recursions are based on the filtering estimates.

5.5.1 Least squares priors

In the literature it is common to use LS estimations to fix the prior distribution of the states (see e.g. Baumeister, Durinck and Peersman (2008)). Usually the advantage of an LS initialization is that it is less complex than other existing approaches but the trade-off is that there are less leverages to control the prior expectations (compared to more comprehensive approaches). However, in a range of comparable studies this kind of initialization worked well.

In a first step we assume the initial state vector to be normally distributed with conditional expected value (or respectively conditional mean) $\beta_{0|0}$ and conditional variance-covariance matrix $P_{0|0}$. Then the prior state vector is calibrated on the least squares point estimates of a time invariant VAR(p)-process with the same lag order p as for the time varying parameter VAR. This gives us $\beta_{0|0} = b_0^{LS}$ as the prior mean. The corresponding $m \times m$ diagonal LS prior variance-covariance matrix is calibrated in a similar way that allows for an additional regulating leverage called τ . The variance of each prior state is simply set equal to the corresponding main diagonal element of the LS estimated variance-covariance matrix $\hat{\Sigma}^{LS}$ for the constant parameter VAR(p). All the initial covariances are set to zero as we have assumed no correlation between the state residuals in our state space model assumptions. In a last step, the elements of this matrix are multiplied by the parameter τ , which postulates the analyst's confidence that the LS prior expectations for the constant terms and lag coefficients in the state vector are binding for the posterior states. For higher values of τ the prior states become less binding for the posterior states and thus the analyst's confidence in the priors is lower. Therefore we call this parameter the “**tightness on the state coefficients**”. Summing it up, the initial state vector is now normally distributed with conditional expected value b_0^{LS} and conditional variance-covariance matrix $P_0^{LS} = \tau \cdot \text{diag}(\hat{\Sigma}^{LS})$,³⁰

$$\beta_0 \sim N(b_0^{LS}, P_0^{LS}).$$

Finally and with respect to the chosen estimation procedure, we need to specify the nature of the residual variance-covariance matrix of the state equation, Q . We have to specify Q exogenously as in this version of the paper only the variance covariance matrix of the measurement errors (H) is set to depend on some randomly drawn and normally distributed parameters, PAR .³¹ And thus, only the residual covariance

³⁰Where $\text{diag}(\hat{\Sigma}^{LS})$ is a diagonal $m \times m$ matrix with the LS estimates on the main diagonal.

³¹We further estimated versions where we initially assumed the variance-covariances of the residuals in the state and the measurement equation H and Q , both to be dependent on some randomly chosen and normally distributed parameters. In these more comprehensive versions the optimal residual variance-covariance matrices H^* and Q^* are both estimated by evaluating the likelihood function of the model conditional on the data and given some initially estimated state vectors as

Table 1: Prior parameter value

Parameter	Value
τ	4
λ	10^{-7}

matrix of the measurements is estimated by evaluating the likelihood function of the model conditional on the data and given some initially estimated state vectors as an initial set of hyperparameters to run the procedure. To set the Q we used information from the properties of our normal linear state space model and additionally referred to established specifications in the literature. The state residuals have to be uncorrelated ($E(\eta_t \eta_s') = 0$) and thus only the elements on the main diagonal can be different from zero. Moreover, following Muscatelli et al. (2007) we set Q equal to the prior variance-covariance matrix of the states, in our this initialization scenario this is $P_0^{LS} = \tau \cdot \text{diag}(\hat{\Sigma}^{LS})$, weighted by another controlling leverage λ that is predefined by Doan et al. (1984). Thus, $Q = \lambda \cdot P_0^{LS}$, where λ is called the **tightness on the state residuals**.³²

5.5.2 An alternative approach: Some literature priors

An alternative and a bit more sophisticated initialization approach is the one of Muscatelli, Spinelli and Trecroci (2007) which is based on the findings of Doan, Litterman and Sims (1984). In their analysis of macroeconomic shocks, structural change and real exchange rates Muscatelli et al. (2007) assumed the initial state vector to be normally distributed as well with conditional expected value (or respectively conditional mean) $b_0 = \beta_{0|0}$ and conditional variance-covariance matrix $P_0 = P_{0|0}$,

$$\beta_0 \sim N(b_0, P_0).$$

The b_0 is set to be an initial state vector with “ones” at the coefficient positions corresponding to the own variable at lag 1 and zeros for all other coefficients (constants

initial set of hyperparameters to run the procedure. The main problem of such larger versions is that due to the large number of parameters to estimate the system is always in danger to become overparameterized. With respect to our relatively small sample of around 154 (observations depending on the specification) overparameterization is a serious risk. Indeed the ML estimations lead to rather incredible state coefficients and in many cases to exploding impulse responses at some points of time. In short, the versions where only the measurement covariance matrix is estimated lead to significantly more stable results.

³²We have experimented with specifications for the tightness on the state coefficients τ reaching from 10^{-6} to 10^6 to analyse the effects on the dynamics of the states as described above. The final setting of the tightness on the state coefficients refers to the approach of Baumeister, Durinck and Peersman (2008) who set τ equal to 4 which is frequently done in the bayesian literature.

and lags), respectively for each equation. According to Muscatelli et al. (p. 1413) this prior expectation is chosen such that changes in each of the involved endogenous variable y_i are “difficult to forecast”. This is because only the first own lag of the variable has strong influence on the forecast as its coefficient is near one (1 plus a zero mean disturbance component) and all other coefficients are closely to zero (0 plus a zero mean disturbance component). The corresponding $m \times m$ diagonal prior variance-covariance matrix of the states P_0 is set in a slightly more comprehensive manner, that implies some other useful regulating leverages on the initial states, in addition to the already introduced tightness of the state residuals, λ . However, for some specifications of the leverages we generated results similar to the results based on the LS initialization.³³ Therefore, and to keep the initialization section concise we will only refer to the results from the first procedure. Nonetheless, further work will be done in the field of prior setting.

5.6 Estimation results

5.6.1 Maximum likelihood estimates

Based on our two datasets we have estimated the parameters of the underlying state space models using a maximum likelihood procedure. In our specification of the models only the variances-covariances of the measurement equation have to be estimated as the state covariance matrix is exogenously determined (done in a first step). In a second step the optimal states are obtained using the Kalman filtering recursions.

5.6.2 The optimal variances and covariances

Table 2 shows the ML estimates of the parameters of interest, the variances of the measurement equation for the output gap (\hat{h}_{11}^2) and the fiscal stance (\hat{h}_{22}^2) as well as their covariance (\hat{h}_{12}^2) for both scenarios, the unadjusted and the adjusted data. Furthermore, the values of the model log likelihood and the determinant of the covariance matrices are given. In addition the corresponding least squares estimates for the parameters of the benchmark time invariant VAR are presented.

³³Definitions of the following parameter specifications are made in Appendix C, where a detailed explanation of the initialization and the leverages is presented. According to Doan et al. (1984) we have set $\vartheta = 630$, $\gamma = 0.07$ and $\lambda = 10^{-7}$. Setting $w^2 = 1/74$, the effects of `gdp_gap` shocks on fiscal stance were much closer to zero than they were in any time invariant parameter VAR estimation on sub-samples or in the time varying parameter VAR estimation. This can be due to a far too small weight on the lags of other variables in the explanation of the fiscal stance. In other words, the relative tightness on the lags of other variables is set too high. Therefore we tried up to a 100 times higher weights which lead to results similar to the results based on the LS initialization.

Table 2: ML Estimates of the Residual Variances-Covariances - Measurement Equation; LS Priors

Parameter	Cyclically unadjusted data (1970Q1-2008Q4)		Cyclically adjusted data (1970Q3-2008Q4)	
	ML Estimates (TVP VAR)	LS Estimates (VAR)	ML Estimates (TVP VAR)	LS Estimates (VAR)
\hat{h}_{11}^2	0.3318*10-4	0.3341*10-4	0.3355*10-4	0.3377*10-4
\hat{h}_{22}^2	0.4960*10-4	0.4989*10-4	0.4902*10-4	0.4929*10-4
\hat{h}_{12}^2	0.0527*10-4	0.0533*10-4	0.0038*10-4	0.0373*10-4
Log Likelihood	1114.99	1125.71	1099.28	1109.97
Determinant	1.6181*10-9	1.5341*10-9	1.6443*10-9	1.5568*10-9

The residual variances based on TVP VARs and constant parameter VARs are very close to slightly smaller TVP VAR variances for both datasets. The same holds for the covariances in the unadjusted dataset. In general, this can be taken as evidence for the reliability of the time varying estimations. In our state space model we assumed the residual variance-covariance matrix to be time-invariant as it is the case the benchmark VAR. Therefore, estimates that are close to the benchmark results seem to be speak for the quality of the time varying model. However, as we used the LS estimations as first and second moment of the states prior distribution, this result is not really astonishing. In both model categories the values of the log likelihood functions are slightly higher for the unadjusted data. Comparing both set-ups, the log-likelihood values are a little bit higher in the constant parameter case. However, any statements on the quality of the models have to be based on tests on the likelihood ratios which will be part of future work.

The optimal states

The optimal estimates for the time varying state variables are provided by the final Kalman filter recursion, using the information of the estimated variance-covariance matrix of the measurement disturbances. Figures of the state variables within the corresponding two-standard error bounds are available from the authors upon request.

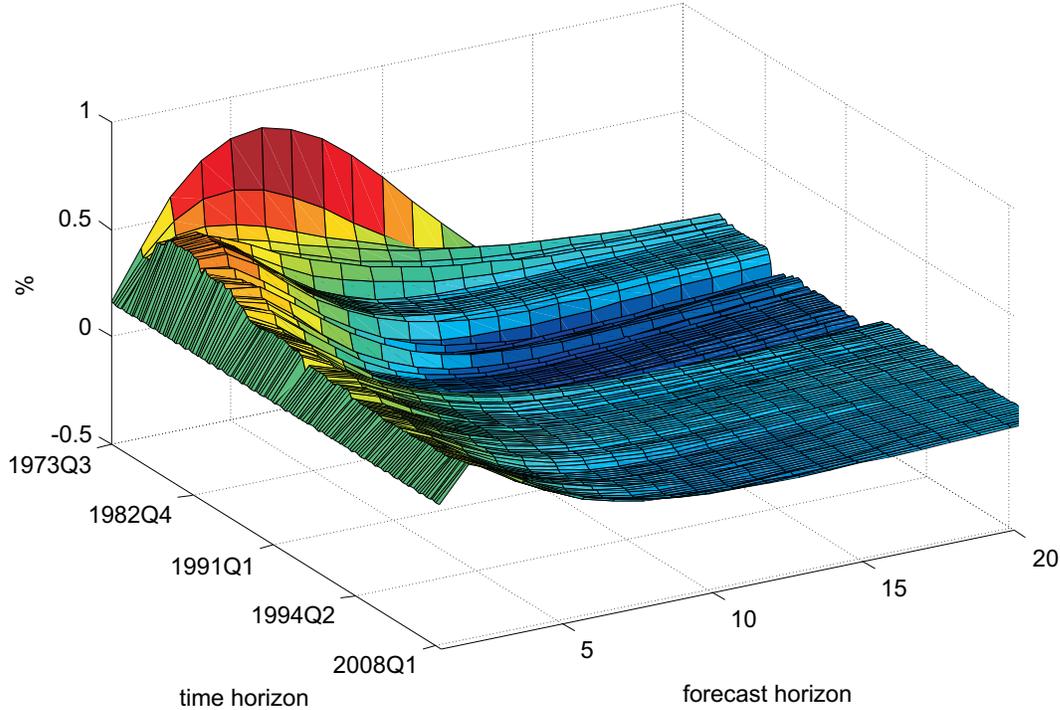
5.7 Structural analysis

5.7.1 Analysis of cyclically unadjusted data in a time-varying framework:

How do the results change if we estimate the effects of the business cycle on fiscal policy within a time-varying framework? We start with a discussion of the results

in our reference Scenario (LS 1 - see appendix) with cyclically undadjusted data represented in figures 3-5. Data on the estimated Variances and Covariances can be found in appendix C

Figure 3: Time-varying parameter SVAR Impulse responses - cyclically unadjusted data
SVAR Orthogonal Impulse Responses of Fiscal Stance to a 1% Output-Gap Shock



In figures 3, 4 and 5 - which reflect the same impulse response functions of the governments' fiscal stance to an output gap shock of one percentage point from different angles - we see first a contemporary reaction of the deficit to changes in the GDP gap, which reflects the working of the automatic stabilizers. The effect equals 0.155 - the same value as in the time-invariant model. Thereafter the surplus increases further reflecting in large parts the further fading-in of automatic stabilizers. Here we can see that the effects have been stronger in earlier times of the period analyzed which might be interpreted as evidence for larger automatic stabilizers in earlier years.

Figure 4: Time-varying parameter SVAR Impulse responses - cyclically unadjusted data
SVAR Orthogonal Impulse Responses of Fiscal Stance to a 1% Output-Gap Shock

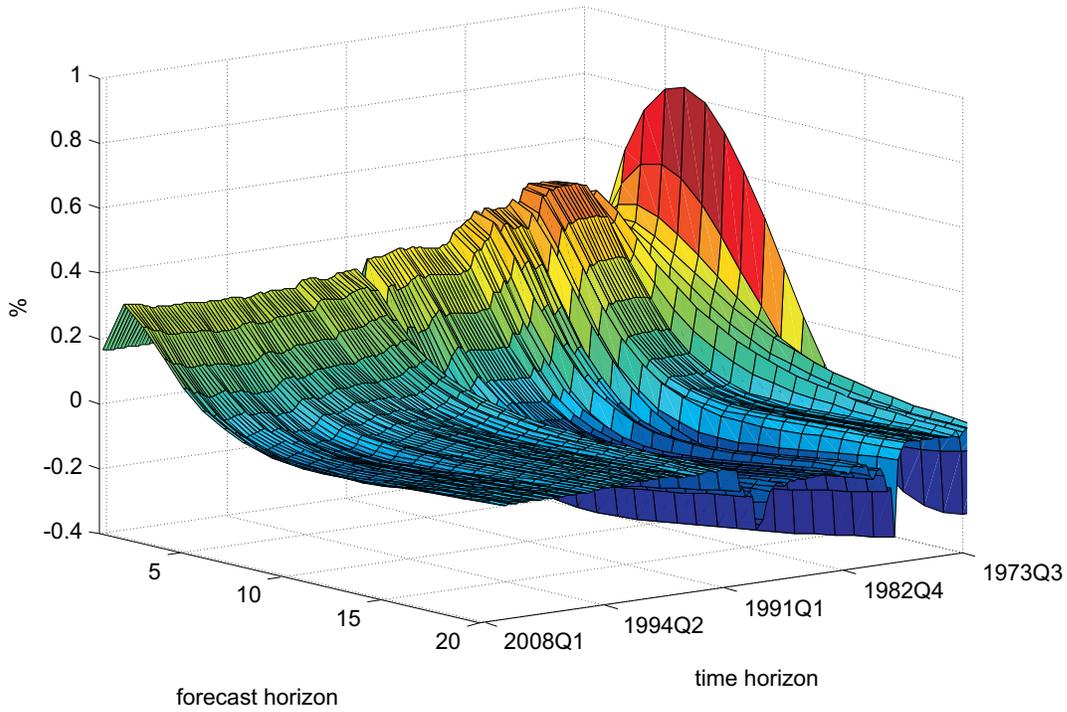
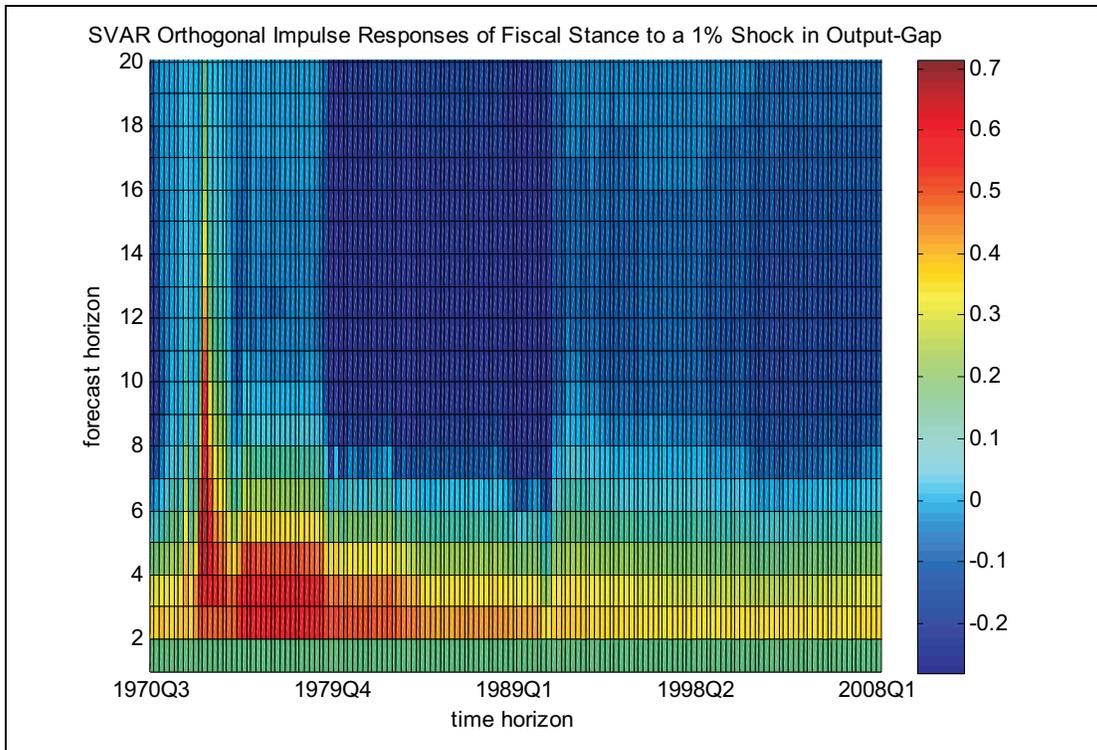


Figure 5: Time-varying parameter SVAR Impulse responses - cyclically unadjusted data

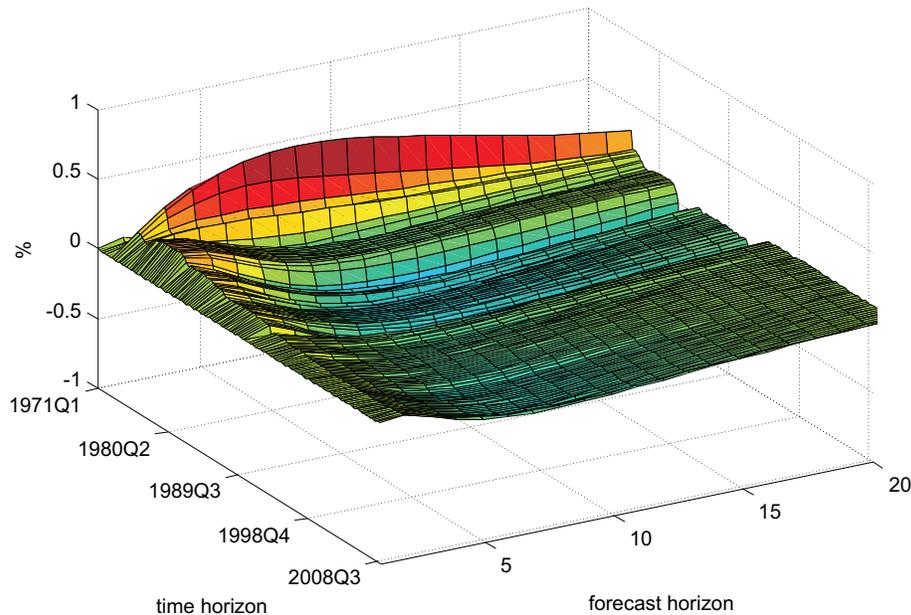


Most important for our analyses however is that we can distinguish three distinctive periods. In the first period from the start of our dataset to 1979 fiscal policy has been strongly counter-cyclical for at least 2 and a half years and never really turned pro-cyclical (see figure 6). After 1979 and until reunification in 1990 however, fiscal policy usually started to be pro-cyclical already after around 6 quarters. As the cyclically unadjusted data reflects the combined effects of discretionary policy and the - always counter-cyclical - automatic stabilizers this points at a strong pro-cyclical timing of fiscal policy from 1979 to 1990 (see as well the analysis of cyclically adjusted data). After 1990 overall fiscal policy - consisting of discretionary measures and automatic stabilizers - became less pro-cyclical again.

5.7.2 Analysis of cyclically adjusted data in a time-varying framework:

In figures 6-8 - which reflect the same impulse response functions of the government's fiscal stance to an output gap shock of one percentage point from different angles - we do not see a contemporary reaction of the deficit to changes in the GDP gap, as these are estimated based on cyclically adjusted data.

Figure 6: Time-varying parameter SVAR Impulse responses - cyclically adjusted data
Forecast Error Impulse Responses of Fiscal Stance to a 1% Shock in Output-Gap



However, the fiscal policy reaction in the first quarter after the shock is in almost all quarters before reunification positive, which might indicate that we have not yet managed sufficiently to take out the working of automatic stabilizers by the method of cyclical adjustment applied. This corresponds as well with our findings in the

time-invariant benchmark model. Besides this we still find evidence for the same three distinct periods from 1970 to 1979, from 1979 to 1990 and afterwards. As we have taken out large parts of the automatic stabilizers now we can get a better feeling of how pro-cyclical fiscal policy has been especially in the period from 1979 to 1990. Here we see (see figure 7) that fiscal policy turned in this period to usually have a pro-cyclical effect already one year after the macroeconomic shock. If we consider that fiscal policy decision making might take a whole year from formulating the idea of a policy change and implementation of this change fiscal policy would have been almost always wrong in these years. After reunification the timing of fiscal policy improved slightly again.

Figure 7: Time-varying parameter SVAR Impulse responses - cyclically adjusted data
Forecast Error Impulse Responses of Fiscal Stance to a 1% Shock in Output-Gap

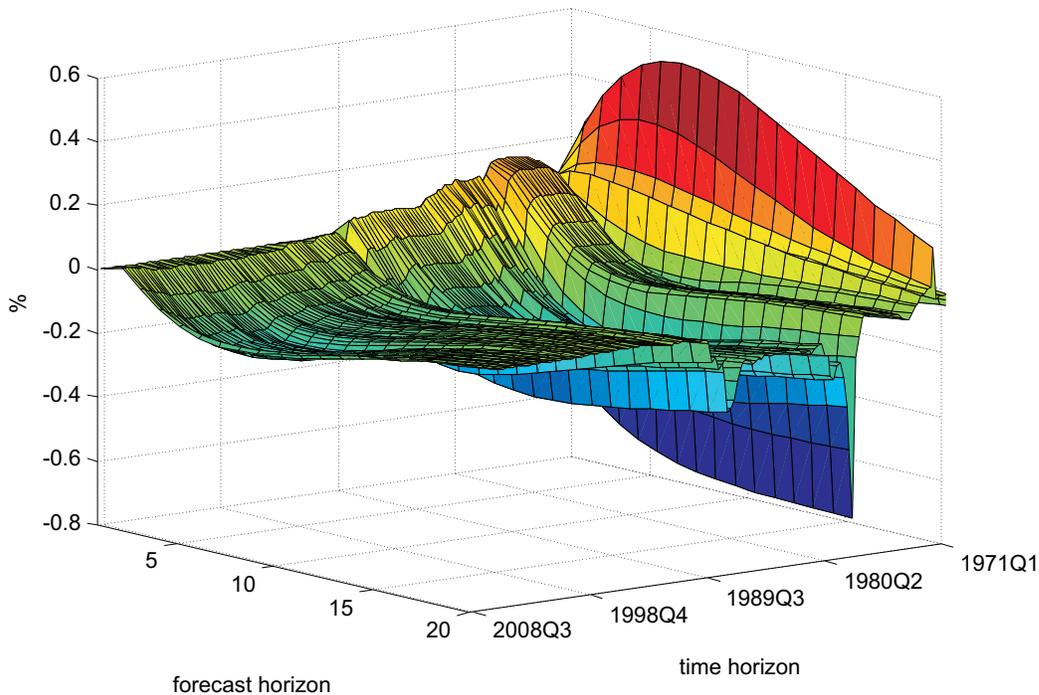
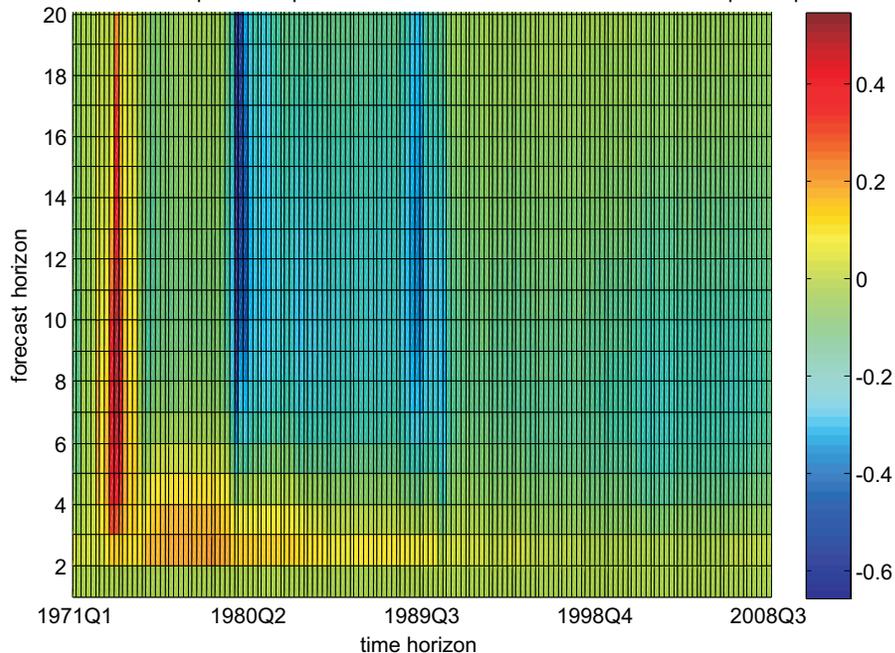


Figure 8: Time-varying parameter SVAR Impulse responses - cyclically adjusted data
Forecast Error Impulse Responses of Fiscal Stance to a 1% Shock in Output-Gap



6 Conclusion and outlook

Large parts of the current empirical economic literature argue that the timing of fiscal policy has been largely a-cyclical in the last decades. Most of these studies rely on cyclically adjusted data on fiscal policy and estimate time-invariant models. For a large dataset for Germany 1970-2008 at quarterly frequency we were able to confirm these results for the relationship of the real GDP gap and the governments fiscal stance in a time-invariant VAR benchmark estimation. However, we found that parameter stability tests strongly point at a parameter instability and therefore a misspecification of the time-invariant model. Based on these findings we derived a time-varying parameter (TVP) VAR, and applied it to the same data. The suspected time-varying regimes with respect to the timing of fiscal policy over the business cycle were strongly confirmed by this approach: we found three distinctive regimes : 1970-1979, 1979-1990 and 1990-2008. During the first regime, policy makers were - from an ex post perspective - successful in timing fiscal policy measures countercyclically. During the second regime, fiscal policy measures have largely been timed pro-cyclically and in the third regime, the timing became slightly more counter-cyclical again. This indicates that the findings by time-invariant approaches of a largely a-cyclical fiscal policy might result from a misspecification of these models, as timing differs strongly across different regimes. Nonetheless there are still many points for an improvement of

the results derived so far. First, the cyclical adjustment procedure applied here seems to be unable to take out all the working of the automatic stabilizers and there is some indication in the results, that automatic stabilizers have been larger in the early years of our data-set. Therefore one possible line for further research is to improve the method of cyclical adjustment - most likely by applying time-varying elasticities within the process of cyclical adjustment. Second, we have estimated VAR models with time-varying parameters but a constant structure of the error covariance over time. Here a robustness check which allows for a varying structure of the error covariance seems to make sense. Furthermore, a more technical point is the calculation and display of median impulse response functions and some impulse responses within 95% Finally a separate analysis of revenue and expenditure developments over the business cycle could help us to understand, which side of the budget is responsible for the timing of fiscal policy over the cycle and could open the door for a more far-reaching research on the causes of different timing patterns in different decades. These points rank high on our research agenda for the next months.

7 Literature

Agnello, Luca and Jacopo Cimadomo (2009): Discretionary Fiscal Policies over the Cycle New Evidence based on the ESCB Disaggregated Approach; *ECB working paper No 1118*/ November 2009.

Alesina A. and Perotti R. (1995), "Fiscal Expansions and Fiscal Adjustments in OECD Countries", *NBER working paper* n. 5214.

Alesina, A., and G. Tabellini (2005) 'Why is fiscal policy often procyclical?', *NBER Working Papers 11600*.

Anderson, B.D.O., and J.B. Moore, 1979: Optimal Filtering. PrenticeHall, 357 pp.

Ballabriga, F. and Martinez-Mongay C. (2002), "Has EMU Shifted Policy?", *Economic Papers of the European Commission, Directorate-General for Economic and Financial Affairs*, No. 166.

Baumeister C. Durinck E. and G. Peersman (2008), "Liquidity, Inflation and Asset Prices in a Time-Varying Framework for the Euro Area" *National Bank of Belgium Working Paper No 142* version

Beetsma, R. and Giuliodori M. (2008), "Fiscal adjustment to cyclical developments in the OECD: An empirical analysis based on real-time data", *CEPR working paper* n.6692.

Bernoth, K., A. Hughes Hallett, and J. Lewis (2008) 'Did fiscal policymakers know what they were doing?', *CEPR Discussion Paper* No. 6758.

Blanchard, O., and R. Perotti (2002), 'An empirical characterization of the dynamic effects of changes in government spending and taxes on output,' *Quarterly Journal of Economics*, 117, 1329-1368.

Blanchard, Olivier Jean and Quah, Danny, 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review, American Economic Association*, vol. 79(4), pages 655-73, September.

Blanchard, Olivier Jean (1990). "Suggestions for a New Set of Fiscal Indicators," *OECD Economics Department Working Papers* 79, OECD, Economics Department.

Bouthevillain, C., Cour-Thimann P., Vand den Dool G., Hernandez de Cos P., Langenus G., Mohr M., Momigliano S., and Tujula M. (2001), "Cyclical adjusted budget balances: an alternative approach", *European Central Bank Working Paper* No. 77, September.

Buti, M. P. van den Noord (2003) 'Discretionary fiscal policy and elections: the experience of the early years of EMU', *OECD Working Paper* , 351.

Caldara , Dario and Christophe Kamps, 2008. "What are the effects of fiscal shocks? A VAR-based comparative analysis," *Working Paper Series* 877, European Central Bank

Chalk N. A. (2002), "Structural Balances and All That: Which Indicators to Use in Assessing Fiscal Policy", *IMF working paper* n.101.

Cimadomo, J. (2007), 'Fiscal policy in real time', *CEPII Working Paper*, No. 10.

Darby, Julia and Jacques Melitz, 2008. "Social spending and automatic stabilizers in the OECD," *Economic Policy*, CEPR, CES, MSH, vol. 23, pages 715-75

Denis, C., K. McMorrow and W. Roeger (2002), 'Production function approach to calculating potential output growth and output gaps - Estimates for the EU Member States and the US', *European Economy*, Economic Paper No. 176.

Doan, T., Litterman, R. B. and C. Smis, 1984. Forecasting and conditional projection using

Eschenbach, F. and L. Schuknecht (2004), 'Deficits and asset prices', *Economic Policy*, July 2004.

Fatás, A. and Mihov, I. (2009), "The euro and fiscal policy", *NBER Working Papers* n.14722.

Fatás, Antonio and Ilian Mihov, 2009. "The Euro and Fiscal Policy," *NBER Working Papers* 14722, National Bureau of Economic Research, In

Fatás, Antonio and Mihov, Ilian, 2001. "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," *CEPR Discussion Papers* 2760, C.E.P.R. Discussion Papers

Forni, L. and Momigliano, S. (2005), "Cyclical Sensitivity Of Fiscal Policies Based On Real-Time Data", *Applied Economics Quarterly* 50(3), 299-326.

Gali, J. and Perotti R. (2003), "Fiscal policy and monetary integration in Europe", *Economic Policy*, 18 (37), p. 533-572.

Gal, Jordi and Roberto Perotti, 2003. "Fiscal policy and monetary integration in Europe," *Economic Policy*, CEPR, CES, MSH, vol. 18(37), pages 533-572, October Gerke (2005)

Giavazzi, Francesco and Marco Pagano, 1990. "Can Severe Fiscal Contractions be Expansionary? Tales of Two Small European Countries," *NBER Working Papers* 3372, National Bureau of Economic Research, Inc.

Girouard, N., and C. André (2005), 'Measuring cyclically-adjusted budget balances for OECD countries', *OECD Economics Department Working Papers* No. 434.

Giuliodori, M. and Beetsma R. (2008), "On the Relationship between Fiscal Plans in the European Union: An Empirical Analysis Based on Real-Time Data", *Journal of Comparative Economics*, vol .2 n. 36, pp. 221-242.

Golinelli, R. and Momigliano S. (2008), "The cyclical response of fiscal policies in the euro area. Why do results of empirical research differ so strongly?", *Bank of Italy, Temi di discussione* n. 654.

Golinelli, R. and Momigliano, S. (2006), "Real-Time Determinants of Fiscal Policies in the Euro Area", *Journal of Policy Modeling*, 28(9), 943-964.

Hallerberg, M. and R. Strauch (2001), 'On the cyclicity of public finances in Europe', *Empirica* 29, 183-207.

Hughes Hallet, A., J. Lewis, and J. Von Hagen, (2004), *Fiscal Policy in Europe, 1991-2003. An Evidence-based Analysis*, CEPR, London.

Jaeger, A. and Schuknecht, L. (2004), 'Boom-bust phases in asset prices and fiscal behaviour', *IMF Working paper* n.54 April 2004.

Jaimovich D. and Panizza U. (2007), "Procyclicality or Reverse causality?", *Working paper series, Inter-american development Bank*, March.

Lane, P., (2003) 'The cyclical behaviour of fiscal policy: evidence from the OECD', *Journal of Public Economics*, 87, 2661-2675.

Larch M. and M. Salto (2005), 'Fiscal rules, inertia and discretionary fiscal policy', *Applied Economics* 37, 1135-1146.

Manasse, P. (2006), 'Pro-Cyclical Fiscal Policy: Shocks, Rules and Institutions - A View from MARS', *IMF Working Paper*, 06 (27).

Moulin, L. and P. Wierts (2006), 'How credible are multi-annual budgetary plans in the EU?', paper presented at the 2006 Banca d' Italia Workshop on Public Finance, Perugia, 30

Muscatelli, Anton V. and Spinelli, Franco and Trecroci, Carmine, 2007. "Macroeconomic shocks, structural change and real exchange rates: Evidence from historical data," *Journal of International Money and Finance*, Elsevier, vol. 26(8), pages 1403-1423, December 2007).

OECD (2003), "Fiscal Stance over the Cycle: Evidence from Panel Analysis", Appendix, Economic Outlook, No. 74. realistic prior distributions. *Econometric Reviews*, 3: 1-144.

Romer C. and Romer D.H. (2007a), "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks.", *NBER working paper series* n. 13264.

Romer C. and Romer D.H. (2007b), "Do Tax Cuts Starve the Beast? The Effect of Tax Changes on Government Spending ", *NBER working paper series* n. 13548.

Romer C. and Romer D.H. (2008), "A Narrative analysis of post-war tax changes", *Unpublished paper*, University of California, Berkeley April 2008.

Turrini A. (2008), "Fiscal Policy and the cycle in the Euro Area: The role of government revenue and expenditure", *European Economy, Economic Papers* 323, May 2008.

Wyplosz, C. (2006), "European Monetary Union: The Dark Sides of a Major Success", *Economic Policy*, April, pp. 208-261.

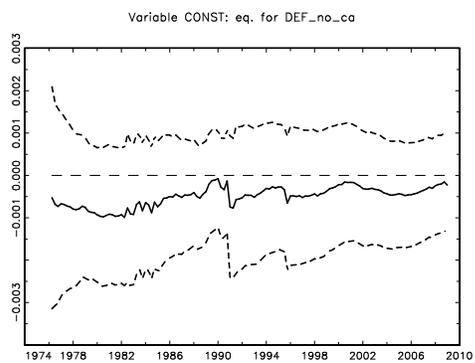
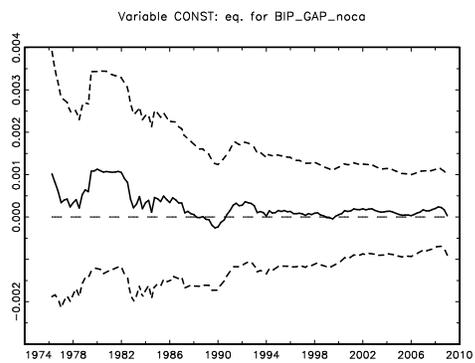
8 Appendices

8.1 Appendix A:

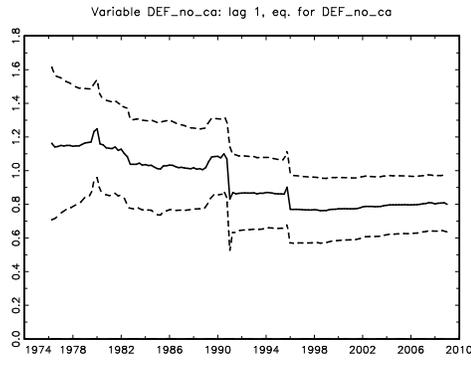
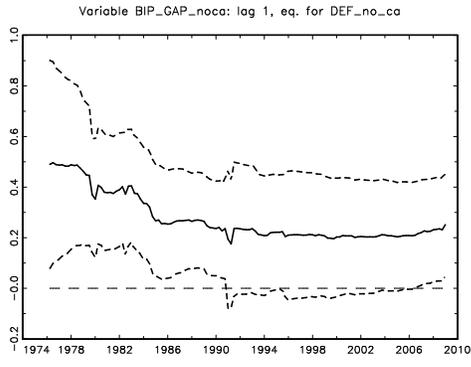
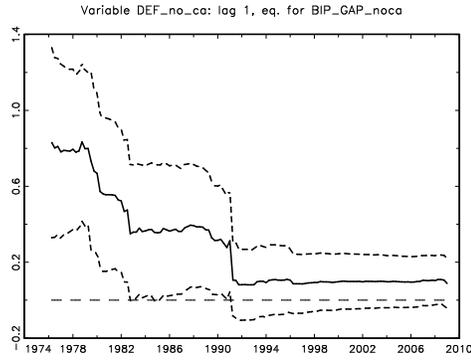
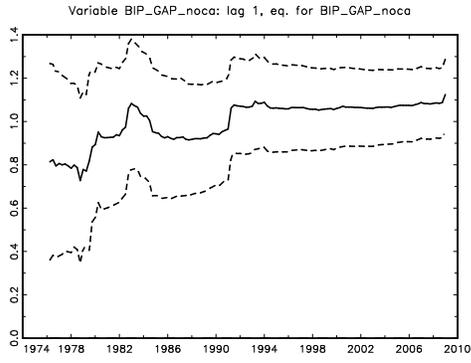
8.1.1 Appendix A:i Stability Analysis for time-invariant benchmark model (Choleski identification): UNADJUSTED DATA

Method: Recursive coefficients estimated by multivariate LS.

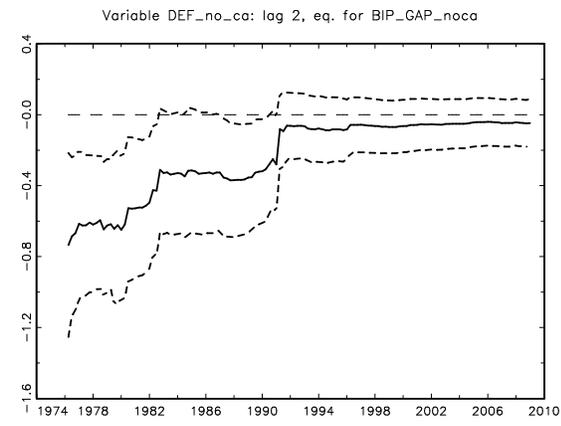
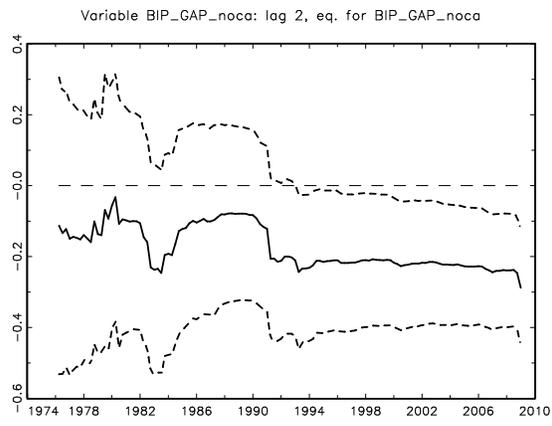
1. Constants (gdp_gap/def):

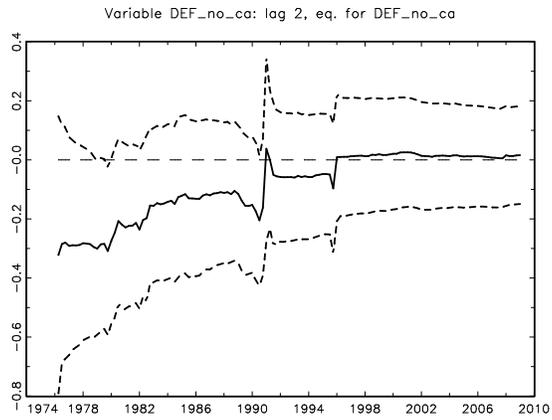
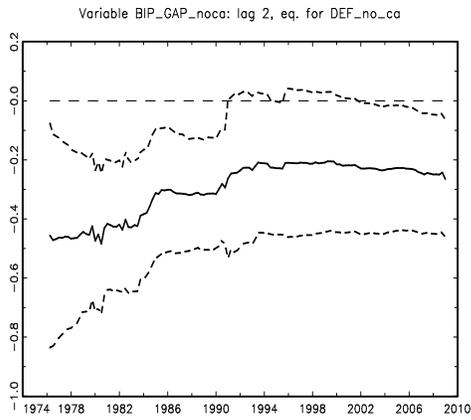


1. Lag 1 (gdp_gap/def):



1. Lag 2 (gdp_gap/def):



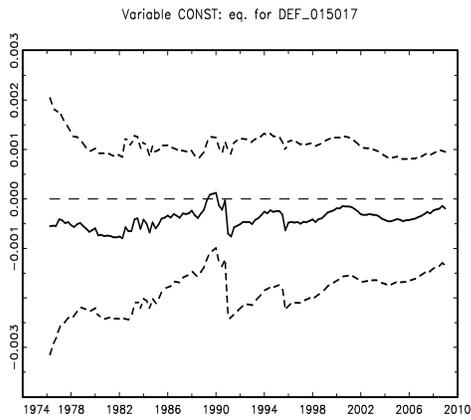
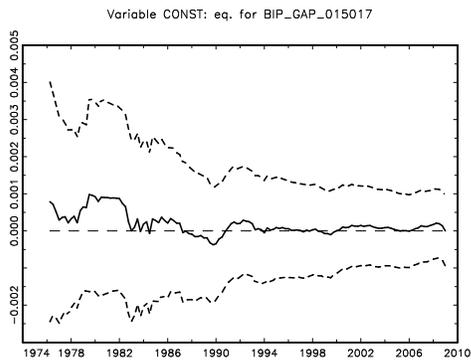


Results: indicating parameter instability

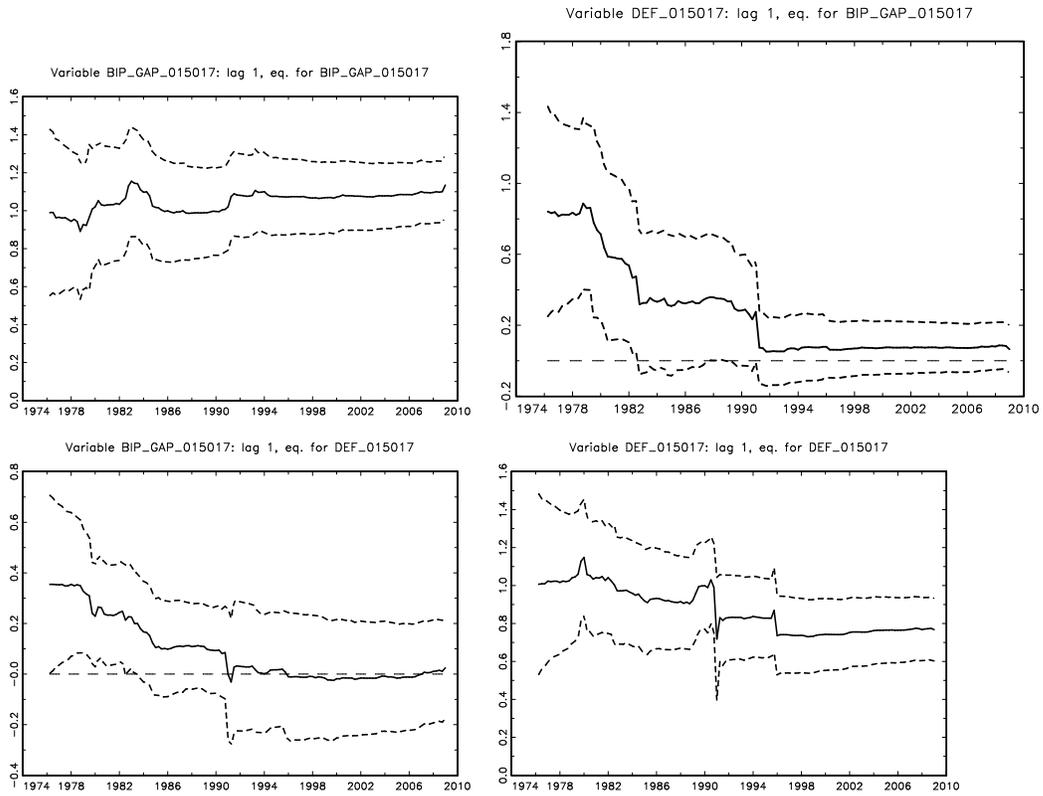
8.1.2 Appendix A:ii Stability Analysis for time-invariant benchmark model (Choleski identification): ADJUSTED DATA

Method: Recursive coefficients estimated by multivariate LS.

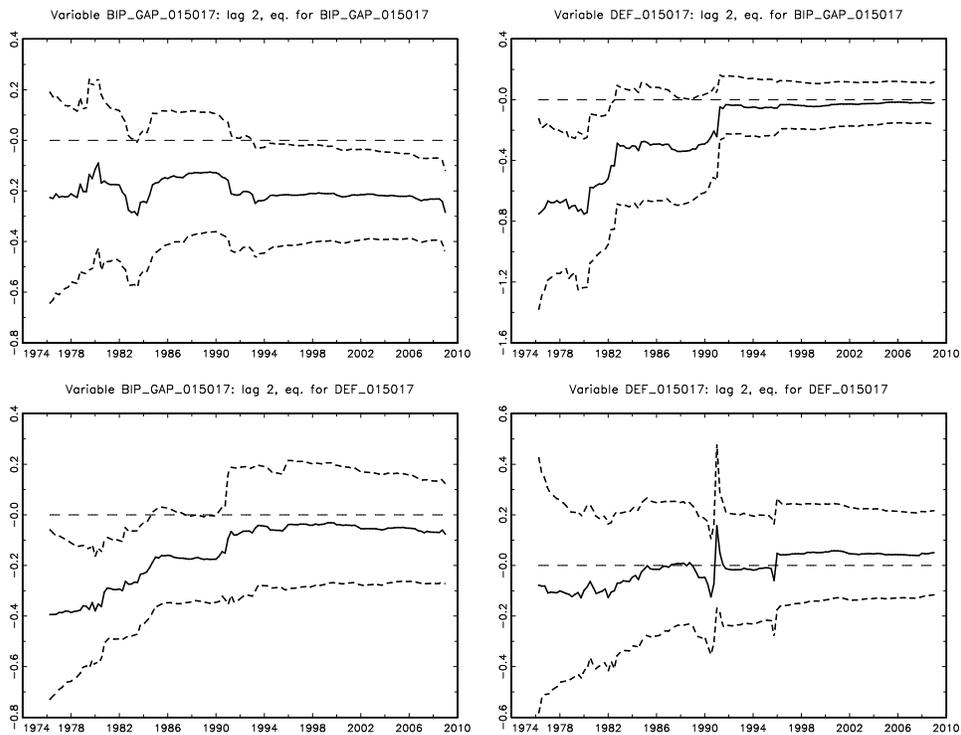
1. Constants (gdp_gap(1st row) / def (2nd row):



1. Lag 1 (gdp_gap(1st row) / def (2nd row):

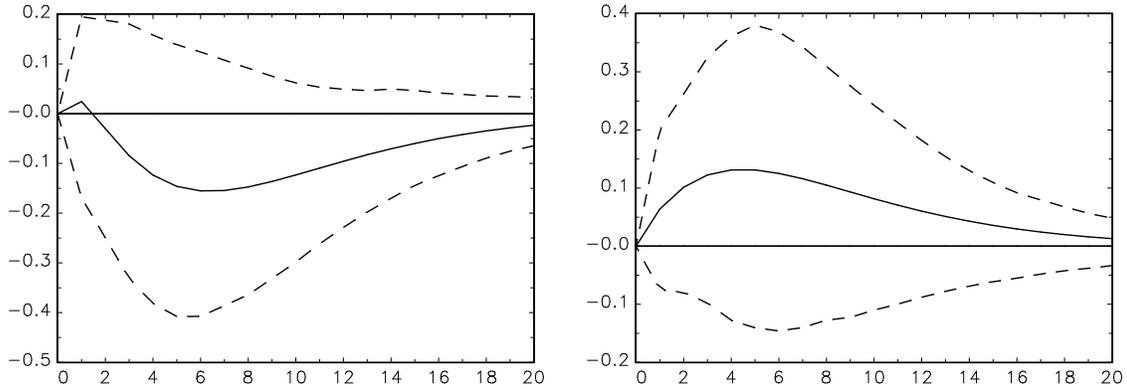


1. Lag 2 (gdp_gap(1st row) / def (2nd row):



1. Results: serious problems with parameter stability

8.1.3 Appendix A:iii Forecast-Error-Responses:



8.2 APPENDIX B:

8.2.1 B.1 The log-likelihood function

The log-likelihood function of our Gaussian model can be derived in three steps.³⁴ In a first step, using Bayes' theorem and the sample density function, the joint density function can be derived as follows,

$$\begin{aligned}
 L(y; H, Q, b_0, P_0) &= f(y_1, \dots, y_T; H, Q, b_0, P_0) \\
 &= f(y_1; H, Q, b_0, P_0) \cdot f(y_2, \dots, y_T | y_1; H, Q, b_0, P_0) \\
 &\quad \vdots \\
 &= f(y_1; H, Q, b_0, P_0) \cdot f(y_2 | y_1; H, Q, b_0, P_0) \cdot \dots \cdot f(y_T | y_1, \dots, y_{T-1}; H, Q, b_0, P_0) \\
 &= \prod_{t=1}^T f(y_t | Y_{t-1}; H, Q, b_0, P_0) \\
 &= \prod_{t=1}^T f(y_t | Y_{t-1}; \theta)
 \end{aligned}$$

Where $\theta = (H, Q, b_0, P_0)$ is a vector of hyperparameters and $f = f(y_t | Y_{t-1}; \theta)$ the distribution of y_t , conditional on the information set at time $t - 1$, Y_{t-1} . Based on this information set, Y_{t-1} , and given a specific set of hyperparameters, the true

³⁴The derivation is based on Ltkepohl (2006, p. 632)).

state vector of our model at time t is by assumption normally distributed with mean b_t and variance-covariance matrix P_t . Therefore, also y_t is normally distributed with recursively given mean $E(y_t | Y_{t-1}) = y_{t|t-1}$ and variance-covariance matrix $Cov(y_t | Y_{t-1}) = F_t$, where the Kalman filter recursions can be used to estimate these quantities, given a specific vector θ . Thus, $y_{t|t-1} = Zb_{t|t-1} + d$ and $F_t = ZP_{t|t-1}Z' + H$.³⁵ This leads to the following log-likelihood function that has to be evaluated for our Gaussian model, where the dimension of our data matrix is denoted by a general dimension n and ν_t is the estimation error from the Kalman filtering procedure,

$$\ln L(\theta | y) = \ln L(H, Q, b_0, P_0 | y) = -\frac{nT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |F| - \frac{1}{2} \sum_{t=1}^T \nu_t' F_t^{-1} \nu_t$$

8.3 APPENDIX C:

8.3.1 APPENDIX C. 1. An alternative initialization approach

The prior mean of the states

In the vectorised form of the state-space model where the measurement equation is described by an n -dimensional TVP-VAR of order p and an $m = (n + pn^2)$ -dimensional state vector, this would then lead to the following vector of prior expectations,

$$b_0 = \left(\begin{array}{c} \overbrace{\left(\begin{array}{cccc} 0 & 0 & \dots & 0 \end{array} \right)}^{\text{constant}} \quad \overbrace{\left(\left(\begin{array}{cccc} 1 & 0 & \dots & 0 \end{array} \right) \left(\begin{array}{cccc} 0 & 1 & \dots & 0 \end{array} \right) \dots \left(\begin{array}{cccc} 0 & 0 & \dots & 1 \end{array} \right) \right)}^{\text{lag 1}} \quad \dots \\ \underbrace{\hspace{15em}}_{(1 \times n^2)} \\ \dots \quad \overbrace{\left(\left(\begin{array}{cccc} 0 & 0 & \dots & 0 \end{array} \right) \left(\begin{array}{cccc} 0 & 0 & \dots & 0 \end{array} \right) \dots \left(\begin{array}{cccc} 0 & 0 & \dots & 0 \end{array} \right) \right)}^{\text{lag p}} \\ \underbrace{\hspace{15em}}_{(1 \times n^2)} \end{array} \right)' \quad (m \times 1)$$

The prior variance-covariance of the states

For all the states (constant terms and lag coefficients) that can be found in the equation for only one variable $i = 1$ the diagonal states' variance-covariance matrix P_0^i is given by

$$P_0^1 = \begin{pmatrix} \vartheta \hat{\sigma}_1^2 & 0' \\ 0 & (G \otimes C^1) \end{pmatrix}_{\substack{(1 \times 1) & (1 \times pn) \\ (pn \times 1) & (pn \times pn)}} \quad (1+pn \times 1+pn)$$

³⁵In our case the vector d is equal to zero.

where

$$G = \begin{pmatrix} \gamma^2/1 & 0 & \dots & 0 \\ 0 & \gamma^2/2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma^2/p \end{pmatrix}_{(p \times p)} \quad \text{and } C^1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & w^2 (\hat{\sigma}_1^2/\hat{\sigma}_2^2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w^2 (\hat{\sigma}_1^2/\hat{\sigma}_n^2) \end{pmatrix}_{(n \times n)}$$

This structure can be transferred to the vectorised form of the state space model for all the n variables:³⁶

$$P_0 = \begin{pmatrix} \vartheta \hat{\sigma}_{(n \times n)}^2 & 0'_{(n \times pn^2)} \\ 0_{(pn^2 \times n)} & (G \otimes C)_{(pn^2 \times pn^2)} \end{pmatrix}_{(n+pn^2 \times n+pn^2)},$$

where

$$\hat{\sigma}^2 = \begin{pmatrix} \hat{\sigma}_1^2 & 0 & \dots & 0 \\ 0 & \hat{\sigma}_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\sigma}_n^2 \end{pmatrix}_{(n \times n)} \quad ; G = \begin{pmatrix} \gamma^2/1 & 0 & \dots & 0 \\ 0 & \gamma^2/2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma^2/p \end{pmatrix}_{(p \times p)}$$

and the elements of the main diagonal of the $n^2 \times n^2$ matrix C written in a $1 \times n^2$ vector ς , are the following,³⁷

$$\varsigma = \begin{pmatrix} \begin{array}{c} \text{Variances of the impacts of all } n \text{ variables in the 1. variable} \\ \text{for only one lag } j = 1, \dots, p \\ \hline 1 \quad w^2 (\hat{\sigma}_2^2/\hat{\sigma}_1^2) \quad \dots \quad w^2 (\hat{\sigma}_n^2/\hat{\sigma}_1^2) \\ \hline (1 \times n) \end{array} & \begin{array}{c} \text{Variances of the impacts of all } n \text{ variables in the 2. variable} \\ \text{for only one lag } j = 1, \dots, p \\ \hline w^2 (\hat{\sigma}_1^2/\hat{\sigma}_2^2) \quad 1 \quad \dots \quad w^2 (\hat{\sigma}_n^2/\hat{\sigma}_2^2) \\ \hline (1 \times n) \end{array} \\ \dots & \dots \\ \dots & \begin{array}{c} w^2 (\hat{\sigma}_1^2/\hat{\sigma}_n^2) \quad w^2 (\hat{\sigma}_2^2/\hat{\sigma}_n^2) \quad \dots \quad 1 \\ \hline \text{Variances of the impacts of all } n \text{ variables in the } n - \text{th variable} \\ \text{for only one lag } j = 1, \dots, p \\ \hline (1 \times n) \end{array} \end{pmatrix}_{(1 \times n^2)}$$

The term $\hat{\sigma}_i^2$ is the estimated residual variance that comes from regressions for univariate AR(p)-processes for each of the $i = 1, \dots, n$ series separately.³⁸ Corresponding

³⁶Thereby one has to pay attention to the fact the ordering of the state coefficients in the vectorised form differs from the one in matrix notation which. The ordering in the state vector – starting with the coefficients of the constant of each equation and continued by the coefficients that model the impact of the first variable on each of the other variables in the first lag, and then the coefficients that model the impact of the second variable on each of the other variables in the first lag and so on until the impact of the n -th variable in the first lag is modelled, continued until the p -th lag – equals the ordering of the variances that can be found on the main diagonal of the variance-covariance matrix of the states. The covariances are generally assumed to be zero.

³⁷The complete matrix is presented in B.2.

³⁸The lag order of the univariate AR(p) processes corresponds to the lag order chosen for the TVP-VAR

to the order of the coefficients the $n \times n$ matrix $\hat{\sigma}^2$ includes the variances of the constants, whereas the variance of the constant in equation i is set to the residual variance of series i . All variances are weighted by the factor ϑ , that can be used as a leverage on the prior expectation for the constants in b_0 where the corresponding terms are set to zero. Thus, the higher the ϑ is set, the less binding is the initial zero restriction on the constants. According to Doan et al. we call this factor the **“relative tightness on the constant terms”**.

The factor γ is called the **“relative tightness on the lag coefficients”** and might be used to regulate how binding the prior expectations on each lag coefficient in b_0 are, thereby including the coefficient of a variable at its own first lag which is restricted to unity in each equation. The higher γ^2 (γ), the higher the variances of all the impact-coefficients in each variable in e.g. one specific lag $j = 1, \dots, p$, and thus the less binding the prior expectation on all of these coefficients is, whereof most are set to zero. However, this also includes the prior expectation on the impact of the own variable at lag 1 that is set to 1 for each series.³⁹

Based on this the $p \times p$ matrix G is used to put a lag-specific “relative tightness on the lag coefficients”. Multiplying all elements of the main diagonal of C that correspond to only lag j by the lag-specific factor γ^2/j , puts a relative weight on the variances of all the coefficients of a specific lag, whereas lags of higher order have less weight than lags of lower order. Therefore the prior expectations on higher lag order coefficients are relatively more binding as variances in higher lags are relatively smaller. As the prior expectations on these coefficients are zero, a tighter restriction on them ensures that they are less useful in forecasts than lower lags are. Thus, γ^2/j is called the **“relative looseness on the lags”**.

The elements of the main diagonal of the $n^2 \times n^2$ matrix C are the lag-unweighted variances of the impacts of variable 1 in each of the n equations in the first lag followed by the impacts of the second variable in each of the n equations in the first lag and so on, written one below the other. The lag-unweighted variances of the own lags of each variable i are set to one, whereas the other elements are ratios of the estimated variances from univariate regressions for each series. These elements are further multiplied by another factor w^2 that additionally weights the impacts of the lags of all other variables on e.g. the variable i . The smaller the w^2 , the more binding the zero expectation from the priors in b_0 on the corresponding coefficients and the less useful they are in the forecast of the variable i . Thus, this leverage is called the **“relative tightness on lags of other variables”**.

³⁹As the variances of each variables own first lag coefficient that is set to unity in b_0 are set to 1 before being weighted by γ^2 , the weighted variances of these first lag coefficients are just $\gamma^2 \cdot 1$. Thus, according to Muscatelli et al. (p. 1414) the γ that is then the standard deviation of the first lag coefficient can be defined as the “analyst’s confidence that the first-order autoregressive coefficients is near to unity”. The lower the standard deviation, the higher the confidence.

Finally, and according to the Muscatelli et al. scenario we need to specify the nature of the residual variance-covariance matrices of the state space model representation. Having defined the prior expectations and their variances we follow the authors in setting the diagonal variance-covariance matrix of the residuals of the state equation, Q , equal to the prior variance-covariance matrix of the states multiplied by a fixed factor that is suggested by Doan et al. (1984). We set $Q = \lambda \cdot P_0$ and call λ the “**tightness on the state residuals**”. Thus, Q does not depend on some randomly chosen normally distributed parameters, PAR , whereas the variance-covariance matrix of the measurement errors, H , does.

Summing it up, the alternative initialization scenario based on the approach of Muscatelli et al. (2007) offers at least five leverages to control and manage the prior setting process which is useful at least with respect to a possibly strong impact of the priors on the estimation process. Moreover, the literature provides predefined values for this initialization approach that have worked well in other long time series analyses. Doan et al. (1984) suggest these predefined values to be:

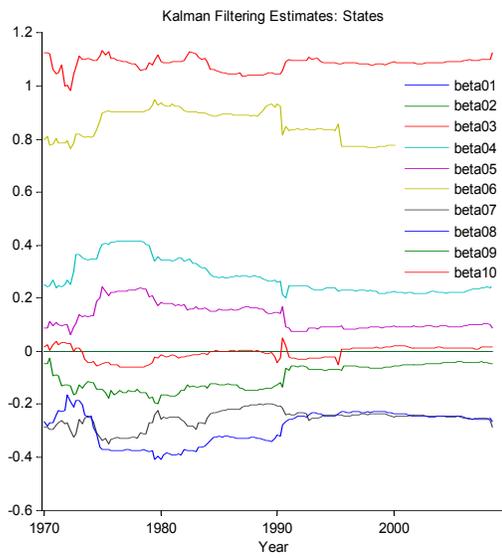
Parameter	Value
ϑ	630
γ	0.07
w^2	1/74
λ	10^{-7}

8.4 Appendix D: Estimation Results

i) Cyclically unadjusted Data (start: 1970Q1)

i.2) Estimated States

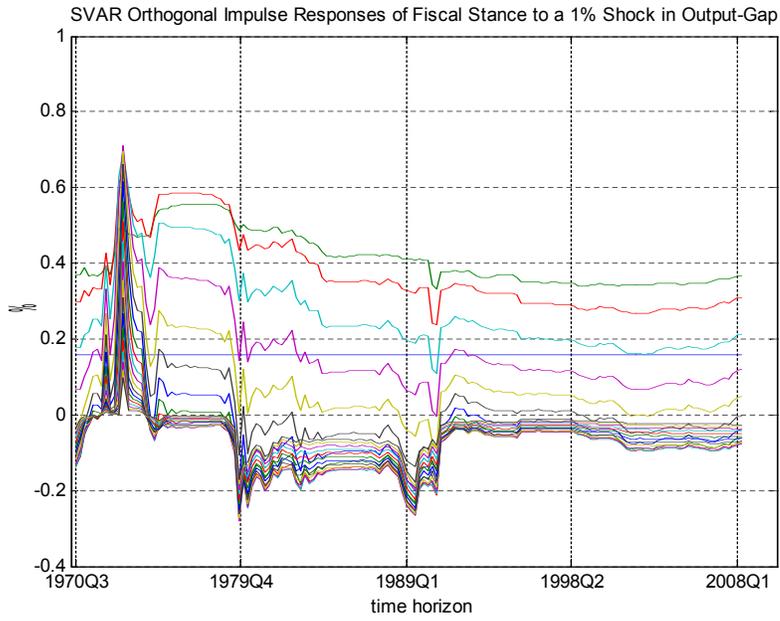
1. (a) **LS-Scenario(1):**



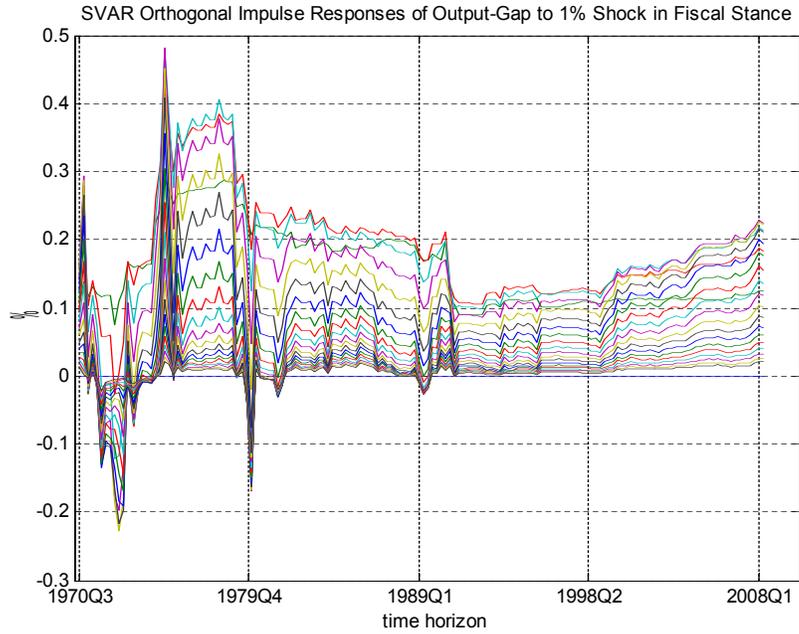
i.3 Impulse Responses

b) LS(1)-Scenario:

GDP_Gap Fiscal Stance



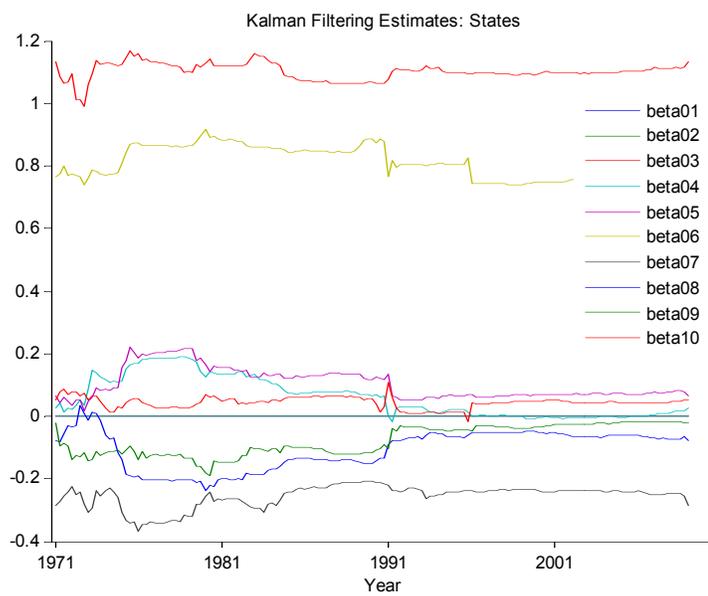
Feedback: Fiscal Stance GDP_Gap:



i) Cyclically adjusted data: (start: 1970Q3)

ii.2 Estimated States

1. (a) LS-Scenario(1):



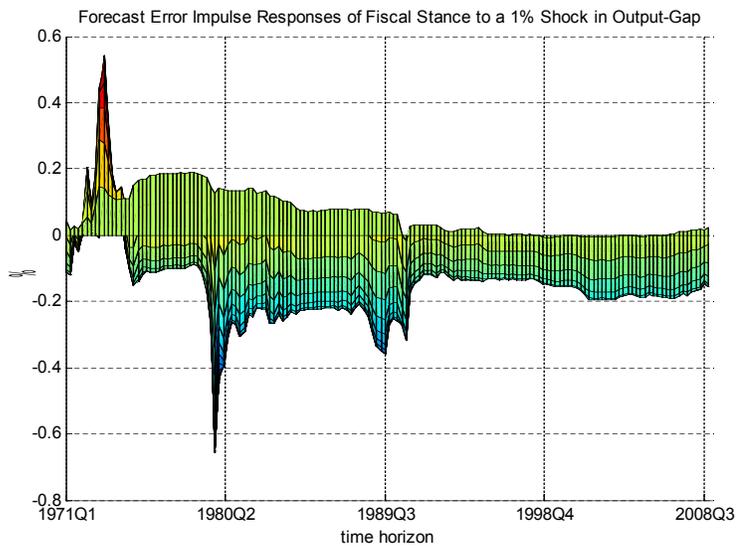
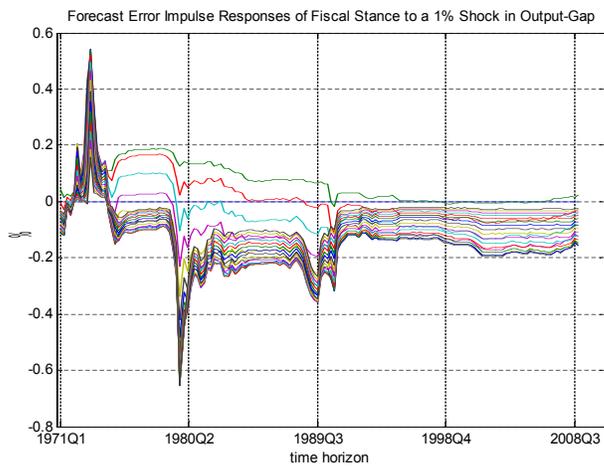
ii.3 Impulse Responses

Start: DATA: 1970Q1 ; VAR: 1970Q3:

b) LS(1)-Scenario:

Forecast Error Impulse Responses

GDP_Gap Fiscal Stance:



Feedback: Fiscal Stance GDP_Gap