Fiscal policy and the business cycle: the impact of government expenditures, public debt, and sovereign risk on macroeconomic fluctuations
Kirchner, M.K.

Citation for published version (APA):
Chapter 2

Transmission of Government Spending Shocks in the Euro Area: Time Variation and Driving Forces*

Abstract

This chapter applies structural vector autoregressions with time-varying parameters to investigate changes in the effects of government spending shocks in the euro area and it studies the driving forces of those changes. We first present evidence on the effects of government spending shocks on real GDP and other variables for individual quarters during the period 1980-2008. We then exploit state dependency using regression inference to add additional structure to the results. Our findings show that short-run spending multipliers have increased from the early 1980s to the late 1980s but they have decreased thereafter until the late 2000s. Moreover, the longer-term effects of spending shocks have declined substantially over this period. We also find that the time variation in spending multipliers can be traced back to increasing availability of credit and rising debt-to-GDP ratios, as well as a smaller share of government investment and a larger share of public wages in total spending.

2.1 Introduction

Fiscal policy has been rediscovered as a tool for short-run economic stabilization. In the context of the recent financial and economic crisis, governments around the world have enacted unprecedented fiscal stimulus packages to counter the severe economic downturn. For instance, the fiscal stimulus adopted within the European Economic

*This chapter is based on joint work with Jacopo Cimadomo and Sebastian Hauptmeier.
Recovery Plan is expected to reach about 1% of the EU’s GDP in 2009 and 0.9% in 2010, and it is largely based on government spending (see European Commission, 2009). However, there is a high degree of uncertainty concerning the macroeconomic impact of government expenditure policies. The theoretical and empirical literature on the effects of government spending shocks reflects this uncertainty as it is rather inconclusive so far, especially as regards the euro area.

Against this background, this chapter offers two contributions. First, we uncover changes in the effects of government spending shocks in the euro area over the period 1980-2008 using the tools of time-varying parameters VAR (TVP-VAR) analysis, allowing for drifting coefficients and stochastic volatility in the VAR model. Second, we provide empirical evidence on the driving forces of the detected time variation in spending multipliers. In particular, using regression inference we relate the estimated multipliers to a set of macroeconomic indicators and to the composition of spending. The underlying idea is that this type of analysis can add additional structure to the results, in a way that may reveal useful information on the fiscal transmission mechanism. To our knowledge, this is the first study that investigates time variation in the effects of government spending shocks through an application of TVP-VAR techniques.\(^1\) In addition, the present study represents the first attempt to provide empirical evidence, by means of a systematic exploitation of state dependency, on the driving factors behind the changing patterns of spending multipliers.

Our analysis is based on a quarterly fiscal data set for the euro area developed by Paredes, Pedregal, and Pérez (2009) for the period 1980-2008. The focus on the aggregate euro area has several advantages. In particular, sub-sample instability has been an obvious possibility at the euro area level, given significant structural changes experienced since the 1980s. Examples include the adoption of the Maastricht Treaty in 1992, the run-up to the EMU, the introduction of the single currency, and the single monetary policy since 1999. Such events should enhance the scope for time variation and help the identification of the key elements of the fiscal transmission mechanism.

\(^1\)TVP-VAR models have been applied to study changes in the effects of monetary policy and the relation to the “Great Moderation” (see e.g. Benati and Murtaz, 2007; Canova and Gambetti, 2009; Cogley, Primiceri, and Sargent, 2010; Cogley and Sargent, 2002, 2005; Galí and Gambetti, 2009, Primiceri, 2005), and the implications of structural change for macroeconomic forecasts (see D’Agostino, Gambetti, and Giannone, 2009).
In addition, while an investigation of time variation at the country level would also be of interest, such an analysis would suffer from the lack of fiscal data sets for single euro area countries of sufficient quality and length.\(^2\)

On the other hand, the use of aggregate euro area data also poses the question of how fiscal shocks should be interpreted in this context. In particular, while the single monetary policy has been in place since 1999 and national monetary policies were largely synchronized before that date, fiscal policy, despite a higher degree of coordination within the EU fiscal framework, remains mainly a country-specific matter. At the same time, the use of aggregate data is justified for the following reasons. First, there is evidence that discretionary fiscal policies have co-moved significantly over the past two decades at the EU and euro area level (see e.g. Giuliodori and Beetsma, 2008). Second, aggregate euro area data can be interpreted as a weighted average of the corresponding country-specific components. This interpretation does not necessarily require that national fiscal policies are aligned. What is instead required is that a spending shock—country-specific or coordinated—is large enough to have an identifiable impact on euro area aggregates. Results are then likely to be driven by shocks occurring in those countries which have the largest weight in euro area variables. Empirical support for this view is provided by Bruneau and Bandt (2003), who show that euro area fiscal shocks were largely induced by Germany, especially in the 1990s. Against this background, a growing number of studies, based on calibrated or estimated DSGE models, now postulates an aggregate fiscal policy for the euro area.\(^3\)

Based on a fixed parameters VAR model estimated over the 1980-2008 sample, our first set of results indicates that, on average, government spending shocks have had an expansionary short-run impact and moderately contractionary longer-term effects on output and the components of domestic private demand in the euro area. However, our time-varying approach uncovers important changes in the effects of spending shocks. In particular, our results suggest that short-run spending multipliers have increased between the early 1980s and the late 1980s but they have decreased thereafter. More-

\(^2\)While the dataset provided by Paredes et al. (2009) is consistent with official national accounts data according to European System of Accounts standards (ESA95), this is not the case for (quarterly) fiscal data of single euro area countries, at least for the period preceding 1999.

\(^3\)See e.g. Christoffel, Coenen, and Warne (2008), Fagan, Henry, and Mestre (2005), Forni, Monteforte, and Sessa (2009), Ratto, Roeger, and in ’t Veld (2009), and Smets and Wouters (2003).
over, the expansionary effects of government spending have become more short-lived over time, and the estimated longer-term effects have decreased substantially. The effects of spending-based fiscal expansions on output indeed appear to be particularly low in the current decade. In addition, smaller spending multipliers on output are found to coincide with a weaker response of private consumption and the real wage, but a stronger response of the short-term nominal interest rate.

With respect to the driving forces of the detected time variation, our evidence points towards availability of household credit as an important determinant of the size of short-run spending multipliers. This result underpins arguments suggesting that access to credit or non-Ricardian behavior by households matter for the effectiveness of fiscal expansions. We also find that a smaller share of investment expenditures and a larger wage component in total government spending are associated with smaller short-run multipliers. Our results therefore seem to provide support for the view that government investment can have positive aggregate supply effects in addition to the aggregate demand effect of government purchases. The fact that wage payments are associated with lower multipliers supports arguments stating that government wage expenditures may have adverse effects in an imperfect labor market through their impact on reservation wages (see Alesina and Ardagna, 2010). Finally, we find that higher debt-to-GDP ratios are associated with lower spending multipliers at longer horizons. This result might suggest that, given higher financing needs of euro area governments, sustained deficits after a spending shock could lead, for instance, to rising concerns on the sustainability of public finances or expectations of a larger future consolidation, which according to our results seems to depress private demand.

The remainder of the chapter is organized as follows. Section 2.2 reviews the related literature. Section 2.3 describes our econometric models, the estimation method, the data, and the structural identification approach. Section 2.4 presents estimation results for the identified models. Section 2.5 discusses several robustness checks. Section 2.6 investigates the driving forces of the detected time variation. It first discusses theoretical views on the fiscal transmission mechanism and, based on this discussion, it identifies the determinants underlying the time variation in spending multipliers using regression analysis. Section 2.7 concludes.
2.2 Related literature

On the theoretical side, there is still considerable disagreement concerning the impact of government spending shocks on important macroeconomic variables. Macroeconomic models used to evaluate the effects of fiscal policy tend to diverge in their predictions (cf. Cogan, Cwik, Taylor, and Wieland, 2010). Neoclassical models with optimizing agents and flexible prices typically indicate a rise in output but a fall in private consumption and real wages following an exogenous increase in government goods purchases (see e.g. Baxter and King, 1993). New Keynesian models, on the other hand, can generate an increase in real wages depending on the monetary regime (see Linnemann and Schabert, 2003). However, basic versions of those models also tend to predict a crowding out of private consumption unless additional features are included which dampen the negative wealth effect of a fiscal expansion. Examples include non-Ricardian consumers (Galí, López-Salido, and Vallés, 2007), imperfect substitutability between public and private consumption (Linnemann and Schabert, 2004), small wealth effects on labor supply (Monacelli and Perotti, 2008), and spending expansions followed by reversals, which create expectations on a future fall in real interest rates (Corsetti, Meier, and Müller, 2009; Corsetti, Kuester, Meier, and Müller, 2010).

On the empirical side, the effects of government spending shocks are typically investigated within the structural VAR (SVAR) framework.\footnote{See e.g. Blanchard and Perotti (2002), Caldara and Kamps (2008), Fatás and Mihov (2001), Mountford and Uhlig (2009), and Perotti (2005).} Alternatives include the event-study approach of Ramey and Shapiro (1998) or, more recently, Ramey (2011b).\footnote{Ramey and Shapiro (1998) and Ramey (2011b) are concerned with the possibility that autonomous fiscal policy changes might be anticipated in advance of their implementation, which is an important challenge for the validity of SVAR results. In this chapter, this issue is addressed in Section 2.5, where we discuss several exercises related to the possible anticipation of the identified SVAR shocks.} Despite an increasing number of studies in this field, many questions remain open. In particular, the effects of government spending shocks in the euro area are largely unexplored. The scarcity of empirical results for the euro area as a whole and also for euro area countries has been mainly due to the lack of quarterly fiscal data, a limitation which has been overcome recently through a quarterly fiscal database for the euro area compiled by Paredes et al. (2009). This data set, which covers the period 1980Q1-2008Q4, is coherent with official annual and quarterly national accounts data, as far as
quarterly fiscal data is available from national accounts (mostly for the period 1999Q1 onwards). Based on this data set, Burriel, de Castro, Garrote, Gordo, Paredes, and Pérez (2010) show that the qualitative responses of macroeconomic variables to fiscal shocks in a (weighted) representative euro area country compare well with results for the U.S. and previous results for some EU countries.

There is also disagreement on whether the effectiveness of fiscal policy has changed over time, and if so to what extent and why. This lack of disagreement concerns especially the effects of government spending, as the literature lacks empirical tests of possible explanations for changing effects of spending shocks. Blanchard and Perotti (2002) have already emphasized that the size of spending multipliers on output in the U.S. varies considerably across sub-periods. Similarly, based on sub-sample or rolling-windows estimation, Bénassy-Quéré and Cimadomo (2006), Bilbiie et al. (2008), Caldara and Kamps (2008), and Perotti (2005) conclude that the responses of the U.S. and of some European economies to fiscal policy shocks have become weaker in the post-1980 period. Perotti (2005) argues that relaxation of credit constraints, a stronger real interest rate response, and changes in monetary policy could explain the decline in the effects of government spending on GDP and its components in OECD countries. Using a New Keynesian model, Bilbiie et al. (2008) show that the more active monetary policy in the Volcker-Greenspan period and increased asset market participation can explain lower spending multipliers in the U.S. after 1980. Overall, confronting potential explanations for changes in the effects of government spending shocks with additional empirical evidence seems a useful contribution to this literature.

2.3 Econometric methodology

Our empirical approach is based on Bayesian estimation techniques. We prefer a Bayesian approach over estimation by classical statistical methods for reasons discussed by Primiceri (2005). Most importantly, a Bayesian approach facilitates the estimation of time variation in multivariate models with drifting coefficients and stochastic volatility. The main advantage of Bayesian techniques is related to the high dimensionality and the non-linearity of such an estimation problem. By using prior information and by
splitting up the original problem into a number of smaller steps, Bayesian methods are able to deal with the high dimension of the parameter space and possible non-linearities in the likelihood function associated with the estimation problem.

We also prefer the TVP-VAR methodology to simpler methods including sub-sample or rolling-windows estimation for the following reasons. First, structural changes could take the form of long-lasting processes, which would not be reflected in an optimal way by sub-sample estimation; they could come suddenly, which would not be reflected by rolling-windows estimation; they could also come suddenly and be reversed afterwards, which would not be reflected in this way by either type of method. Second, structural changes might not be easily identified a priori. Third, one can think of various alternative structural changes which might impact on the effectiveness of fiscal policy, e.g. monetary policy regime changes or trade integration. It would therefore be difficult to date breaks and to determine the size of rolling windows. The TVP-VAR methodology allows to address these issues through estimates for individual quarters.

2.3.1 Reduced-form VAR models

We consider two versions of a reduced-form VAR of lag order \( p \). The first version has fixed parameters:

\[
y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + \Gamma z_t + u_t, \quad t = 1, 2, 3 \ldots, T, \tag{2.1}
\]

where the vector \( y_t \) includes government spending, output, private consumption, the short-term interest rate and possibly other macroeconomic indicators. The \( B_i, i = 1, 2, 3, \ldots, p \), are matrices of coefficients. The vector \( z_t \) collects exogenous variables with parameter loadings \( \Gamma \). The vector of innovations \( u_t \) is assumed to be Gaussian white noise with mean zero and covariance matrix \( R \), i.e. \( u_t \sim N(0, R) \).

The second version generalizes (2.1) by allowing for drifting coefficients and stochastic volatility in the innovations.\(^6\) Both aspects are supposed to capture structural changes such as shifts in private sector behavior and/or changes in the conduct of pol-

\(^6\)Our specification of the TVP-VAR follows Cogley and Sargent (2002, 2005) and Primiceri (2005). We apply some additional restrictions on the hyperparameters as discussed below.
icy. Drifting coefficients are thought to capture changes in the propagation of shocks throughout the economy. Stochastic volatility is introduced to allow for changes in the distribution of the shocks. Hence:

\[ y_t = B_{1,t}y_{t-1} + \cdots + B_{p,t}y_{t-p} + \Gamma_t z_t + u_t, \quad t = 1, 2, 3 \ldots, T, \]  

(2.2)

where \( u_t \sim N(0, R_t) \). Stack the VAR coefficients by equations in a vector \( \beta_t = \text{vec}(F_t') \), where \( F_t = [B_{1,t}, \ldots, B_{p,t}, \Gamma_t] \) and \( \text{vec}(\cdot) \) is the column stacking operator. This state vector of coefficients is assumed to follow a driftless random walk:

\[ \beta_t = \beta_{t-1} + \varepsilon_t, \]  

(2.3)

where \( \varepsilon_t \sim N(0, Q) \). Further, the innovation covariance matrix can be decomposed using a triangular factorization of the form

\[ R_t = A_t^{-1}H_t(A_t^{-1})', \]  

(2.4)

where \( A_t^{-1} \) is lower triangular with ones on the main diagonal and \( H_t \) is diagonal. Stack the elements below the main diagonal of \( A_t \) row-wise in a vector \( \alpha_t \). Collect the diagonal elements of \( H_t \) in a vector \( h_t \). Similarly as the coefficient states, the covariance and volatility states are modeled as (geometric) random walks:

\[ \alpha_t = \alpha_{t-1} + \nu_t, \quad \log h_t = \log h_{t-1} + \omega_t, \]  

(2.5)

where \( \nu_t \sim N(0, S) \) and \( \omega_t \sim N(0, W) \). Thus, following Primiceri (2005), both the diagonal elements and off-diagonal elements of the reduced-form covariance matrix can

---

7 The fixed parameters model (2.1) includes an intercept and a linear-quadratic time trend in \( z_t \) to account for the presence of trends in real variables and the nominal interest rate. A deterministic time trend seems redundant in the TVP-VAR model such that \( z_t \) in model (2.2) includes only an intercept.

8 Compared to alternative specifications such as regime switching models, the random walk specification has the advantage that it allows for smooth shifts in the states of the model. Primiceri (2005) argues that regime switching models may well capture some of the rapid shifts in policy but they seem less suitable for describing changes in private sector behavior, where aggregation usually smoothes out most of the changes, or learning dynamics of both economic agents and policymakers.
drift over time, where the latter allows for changes in the contemporaneous relations among the endogenous variables.

The joint distribution of shocks is postulated as $[u_t, \epsilon_t, \nu_t, \omega_t]' \sim N(0, V_t)$, where $V_t$ is block diagonal with blocks $R_t$, $Q$, $S$, and $W$. Notice that an unrestricted covariance matrix would drastically increase the number of parameters and thus complicate the estimation problem. Independence of $R_t$ and the hyperparameters implies that the innovations to the VAR parameters are uncorrelated with the VAR residuals. This assumption seems plausible since the innovations capture business cycle events, policy shocks, or measurement errors. It seems unlikely that such short-term events are related to longer-term institutional changes and other changes in the structure of the economy, which are captured by the innovations to the VAR parameters. For example, it can be argued that the introduction of the single currency in the euro area has not been related to technology shocks, government spending shocks, and so on.

We make the additional assumption that $Q$, $S$, and $W$ are diagonal to further reduce the dimensionality of the problem and to simplify inference. The assumption of (block) diagonality of $S$ ensures that the rows of $A_t$ evolve independently such that the covariance states can be estimated row by row (cf. Primiceri, 2005). Diagonality of $W$ implies that the volatility states are independent such that the simple univariate algorithm of Jacquier, Polson, and Rossi (1994) can be applied to each element of $u_t$ in order to estimate the volatility states. The reduction of estimated parameters resulting from the diagonality restrictions on $Q$ and $S$ helps to save degrees of freedom in our relatively short euro area data set.

### 2.3.2 Estimation method

Both VAR models described above are estimated by Bayesian methods. For the version with fixed parameters, our prior and posterior for the coefficient matrices $B_i$, $i = 1, \ldots, p$, $\Gamma$, and the covariance matrix $R$ belong to the Normal-Wishart family with a diffuse prior centered on OLS estimates over the full sample. For the TVP-VAR, we apply a variant of the Gibbs sampler (see Geman and Geman, 1984; Smith and Roberts,
The main steps of the estimation algorithm are outlined here whereas Appendix 2.A provides a detailed description. The Gibbs sampler iterates on the following four steps, sampling in each step from lower dimensional conditional posteriors as opposed to the joint posterior of the whole parameter set.

(i) **VAR coefficients.** Conditional on the data and a history of covariance and volatility states, the observation equation (2.2) is linear with Gaussian innovations and a known covariance matrix. The coefficient states $\beta_t$ can thus be sampled using the Kalman filter and a backward recursion, as described in Carter and Kohn (1994) and Cogley and Sargent (2002).

(ii) **Elements of $A_t$.** Conditional on the data and a history of coefficient and volatility states, equation (2.2) can be rewritten as $A_t u_t = v_t$, with $\text{cov}(v_t) = H_t$. This is a linear Gaussian state space system with independent equations, due to the (block) diagonal structure of $S$ (see Primiceri, 2005). The algorithm of Carter and Kohn (1994) can thus be applied equation by equation to sample the elements of $A_t$ on each row below the main diagonal.

(iii) **Elements of $H_t$.** Conditional on the data and a history of coefficient and covariance states, the orthogonalized innovations $v_t$ are observable. Given the diagonal structure of $W$, the diagonal elements of $H_t$ can be sampled using the univariate algorithm of Jacquier, Polson, and Rossi (1994) element by element, following Cogley and Sargent (2005).

(iv) **Hyperparameters.** Conditional on the data and the parameter states, the state innovations $\varepsilon_t$, $\nu_t$, and $\omega_t$ are observable. This allows to draw the hyperparameters (i.e. the elements of $Q$, $S$, and $W$) from their respective distributions.

Under relatively weak regularity conditions (see Roberts and Smith, 1994) and after convergence, iterations on these steps produce a realization from the joint posterior distribution. We generate 60,000 draws from the Gibbs sampler, of which we burn the first 50,000 to let the Markov chain converge to its ergodic distribution. Of the remaining 10,000 draws, we keep every 10th draw to break the autocorrelation of

---

draws. This leaves us with 1,000 draws from the joint posterior distribution of the model parameters. Appendix 2.C analyzes the convergence properties of the Markov chain, concluding that these properties are overall satisfactory.

We follow conventional choices in the calibration of the priors, similar as in Primiceri (2005), but we take a somewhat more conservative stance on the degrees of freedom of the prior distributions which we set to the minimum value allowed for the priors to be proper. Appendix 2.B provides details on the calibration of the priors while the robustness of the results to alternative choices is analyzed in Section 2.5. Unlike most previous studies, we do not truncate the posterior distribution of the VAR coefficients by discarding draws which do not satisfy stationarity conditions. Cogley and Sargent (2002, 2005) have proposed such a restriction for U.S. monetary policy, arguing that the Fed had ruled out unstable paths of inflation. A similar point is harder to defend for aggregate euro area fiscal data since fiscal variables may have followed unstable paths in some countries. We do however check the robustness of the results by imposing stationarity conditions in Section 2.5.

2.3.3 Data description

Our benchmark VAR specification includes government spending, defined as government consumption plus government investment following most of the literature, GDP, private consumption (all in real per capita terms), and the short-term nominal interest rate for the euro area over the period 1980Q1-2008Q4. Real GDP measures economic activity. Private consumption is included since it is the largest component of aggregate demand, and also to be able to contribute to the ongoing discussion on the effects of government spending shocks on that variable. The short-term interest rate is added to this small-scale VAR to assess the impact of government spending shocks on interest rates, and potential changes thereof. We also examine the impact of spending shocks on a broader set of macroeconomic indicators including private investment, net taxes

\[10\] The Gibbs sampler is a dependence chain algorithm. However, independent draws should be used when calculating statistics of interest such as posterior means and impulse responses.

\[11\] Perotti (2005) argues that the long-term interest rate has a closer relation to private consumption and investment decisions than the short-term interest rate. Replacing the short-term interest rate by the long-term interest rate did however not lead to any significant changes in our results.
Chapter 2

Figure 2.1: Data used in the benchmark specification

![Graphs of Gov. Spending over GDP, Private Consumption over GDP, Short-Term Interest Rate](image)

**Notes.** Euro area data, 1980Q1-2008Q4; gov. spending equals final general gov. consumption plus gov. investment; gov. spending and private consumption are expressed as nominal shares of GDP; the short-term interest rate is measured in nominal, annual terms; source of fiscal data: Paredes et al. (2009); source of remaining data: ECB’s Area-Wide Model database.

(i.e. total tax revenues minus transfers), the wage rate, all in real per capita terms, and the annual rate of change of the Harmonized Index of Consumer Prices (HICP).\(^{12}\)

Those additional variables are added all together in the extended specification of the fixed parameters model. In the specification of the model with time-varying parameters we are however constrained by the need to avoid overparameterization and exhausting available degrees of freedom. The additional variables are therefore added one at a time to the benchmark specification, thus limiting the number of variables in the TVP-VAR

\(^{12}\)We use the HICP-based inflation rate to assess the response of inflation to spending shocks due to its close link to monetary policy decisions in the euro area.
to a maximum of five indicators.

As Burriel et al. (2010), we use a quarterly fiscal data set for the euro area compiled by Paredes et al. (2009). The latter employ mixed-frequencies state space models estimated with available (mostly annual) national accounts data and monthly and quarterly fiscal information taken from government cash accounts to obtain interpolated quarterly fiscal data for the above-mentioned period. By construction, the interpolated variables are coherent with official ESA95 annual and quarterly euro area data, as far as the latter is available. This approach has the advantage that it avoids the endogenous bias which could arise if fiscal data interpolated on the basis of general macroeconomic indicators were used with macroeconomic variables to assess the impact of fiscal policies. Other macroeconomic data for the euro area are mainly taken from the ECB’s Area-Wide Model database (see Fagan et al., 2005).

To ensure comparability with the previous literature, our data definitions closely follow related studies. Details are provided in Appendix 2.D. Figure 2.1 shows the data used in the benchmark specification. Both models are estimated in levels and prior to the estimation all variables except the interest rate and the inflation rate were transformed into natural logarithms.

2.3.4 Structural interpretation

The reduced-form models attempt to capture structural representations with uncorrelated shocks. The reduced-form innovations are therefore linear transformations of some underlying structural shocks \( e_t \) with \( E[e_t e'_t] = I \), i.e. \( u_t = C e_t \) for the fixed parameters model and \( u_t = C_t e_t \) for the model with time-varying parameters, for \( t = 1, 2, 3, \ldots, T \). In particular, the innovations in the equation for government spending can be considered as linear combinations of three types of shocks (see Blanchard and Perotti, 2002): (i) the automatic response of spending to movements in the business cycle, prices and interest rates; (ii) the systematic discretionary response of spending to macroeconomic developments; (iii) deliberate discretionary changes in spending. The latter are the truly structural spending shocks of interest.

Without restrictions on the matrices \( C \) and \( C_t \), and therefore the covariance matrices \( R \) and \( R_t \), the above systems are not identified since many combinations of struc-
tural shocks can generate the same reduced-form innovations. To achieve a structural representation, government spending shocks are identified by assuming that government spending is predetermined in a system with output, consumption, the interest rate, and possibly other macroeconomic variables, following Blanchard and Perotti (2002) and Fatás and Mihov (2001). Thus, government spending is ordered first in the estimated models and the desired linear combination is achieved by a Cholesky decomposition, i.e. $R = CC'$ and $R_t = C_tC'_t$, where $C$ and $C_t$ are lower triangular matrices. Under this recursive identification scheme, all variables are allowed to respond within a quarter to innovations to government spending but government spending does not react within a quarter to innovations to other variables in the system.

As discussed by Caldara and Kamps (2008), the fact that government spending as defined here does not include interest payments justifies that spending is ordered before the interest rate. The fact that spending is defined net of transfer payments further justifies the assumption of acyclical, i.e. there is no automatic contemporaneous reaction of spending to movements in the business cycle. In addition, due to implementation lags in the policy process, an immediate discretionary fiscal response to a change in the economy is unlikely to occur. When more variables are included, the assumption that government spending does not react within a quarter to shocks to those variables can be justified on similar grounds.

As mentioned above, a well-known criticism of the above SVAR approach centers on the possibility that autonomous policy changes can be anticipated by economic agents (see e.g. Ramey, 2011b; Leeper, Walker, and Yang, 2011). This criticism is addressed in Section 2.5, based on Granger-causality tests that relate the identified SVAR shocks to institutional forecasts and survey data, following Ramey (2011b).

Impulse responses are then calculated as follows. In the fixed parameters case, given draws from the posterior distributions of $R = CC'$ and the $B_i$, the first column of the matrix $C$ gives the contemporaneous responses (at horizon $k = 0$) of the endogenous variables to a one-time, unitary structural shock to government spending $e_0 = [1, 0, \ldots, 0]'$, and model (2.1) with $u_k = [0, 0, \ldots, 0]'$ can be used to calculate impulse responses at horizons $k \geq 1$. In the time-varying parameters case, we apply a local approximation to the impulse responses at time $t$, following e.g. Galí and Gam-
betti (2009). That is, the contemporaneous responses to unitary shocks \( e_{t,0} \) at time \( t \) are derived from draws from the posterior distribution of reduced-form covariance matrices \( R_t = C_t C_t' \), and the draws from the distribution of the \( B_{i,t} \) are applied to calculate impulse responses at horizons \( k \geq 1 \), using model (2.2) with \( u_{t,k} = 0 \).

2.4 Estimated effects of spending shocks

In this section, we first present estimation results for the identified fixed parameters model, to assess the impact of government spending shocks over the full sample. We then discuss results for the identified time-varying parameters model.

2.4.1 Time-invariant impulse responses

Figure 2.2 reports the estimated impulse responses due to the identified government spending shocks to the four endogenous variables \( y_t \) of equation (2.1) for the benchmark specification, together with their 16 and 84 percent probability bands. The responses of output, consumption, and spending (and later on investment and net taxes) to the spending shock are reported as non-accumulated multipliers. That is, the original impulse responses are divided by the impact response of government spending and the result is divided by the ratio of government spending and the responding variable. The rescaled impulse responses can thus be interpreted to give the reaction of the responding variable, in percent of real GDP, to a spending shock leading to an initial increase in the level of government spending of size 1% of real GDP. For the fixed parameters model, the ratio is evaluated at the sample mean. For the model with time-varying parameters, we take the ratio in the respective quarter.\(^\text{13}\)

According to the results in Figure 2.2, a government spending shock induces a positive, persistent response of spending lasting more than four years. The initial reaction of output is positive, the estimated short-run multiplier being 0.54. The

\(^{13}\)For example, suppose that the shock leads to a two percent increase in government spending. Since the share of spending over GDP is roughly 25 percent, this corresponds to a spending increase of about 0.5% of GDP. Say output increases by one percent and consumption increases by 0.5 percent, i.e. by 0.25% of GDP since the share of consumption over GDP is about 50 percent. The share of spending over consumption is thus roughly 50 percent. The multipliers would be \((1/2)/0.25 = 2\) for output and \((0.5/2)/0.5 = 0.5\) for consumption.
Figure 2.2: Time-invariant impulse responses I – benchmark specification

Notes. Median impulse responses with 16 and 84 percent probability bands; the responses of output, consumption, and spending are measured in percent of GDP to a 1% of GDP spending shock, i.e transformed response at horizon $k = \text{responding variable’s original response at horizon } k/(\text{spending response at horizon } 0 \times \text{average ratio of spending to responding variable});$ the response of the interest rate is measured in percentage points to a one percent shock.

Output response remains positive with 68 percent probability for about one year after the shock, it turns negative after two years, it reaches a minimum after about three years, and it then returns to the baseline. The spending shock also leads to a positive initial response of private consumption. Similarly as for output, however, the response of consumption turns negative over the medium term. The nominal interest rate hardly responds to the spending shock on impact, but it then starts to rise and peaks about one year after the shock. The interest rate response is estimated to be positive with 68 percent probability from two quarters until around three years after the shock.

In a next step we extend the VAR specification by a broader set of indicators
which often appear in related studies. The estimated impulse responses of government spending, output, consumption, investment, the real wage, net taxes, the HICP-based inflation rate, and the nominal interest rate are reported in Figure 2.3. As a consequence of a 1% of GDP spending increase, net taxes increase by about 0.8% of GDP on impact, indicating an overall fiscal expansion since the primary deficit increases. Net taxes also return more quickly to baseline than spending does, thus the shock remains expansionary. Output, consumption, and investment increase at first but fall afterwards below their initial levels. The responses of output and the components of domestic private demand are however estimated with relatively little precision. The
point estimates of the impact multipliers are 0.55 (output), 0.23 (consumption) and 0.03 (investment). The real wage increases by approximately 0.15 percent on impact and remains above its initial level during more than three years after the shock. Inflation shows a muted response in the initial two quarters but it starts to increase later on. The nominal interest rate reacts similarly as in the benchmark specification.

Overall, these results indicate that on average over the period 1980-2008 government spending shocks have had expansionary short-run effects on output, consumption, investment, and real wages in the euro area. However, output declines at longer horizons as consumption and investment are being crowded out. The estimated increase in the nominal interest rate is consistent with an offsetting reaction of monetary policy to the fiscal expansion to reduce inflationary pressures. At the same time, our findings compare well with the results of previous SVAR studies for the euro area. In particular, they are broadly similar to the results of Burriel et al. (2010), the main previous fiscal VAR study for the euro area employing a similar data set. Burriel et al. (2010) also find a positive impact of government spending shocks on GDP and private consumption in the short run and a decline at longer horizons, an increase in the aggregate primary government deficit, and a relatively persistent increase in interest rates.

2.4.2 Time-varying impulse responses

The time-varying nature of model (2.2) allows to compute state-dependent impulse responses for individual quarters of the estimation sample. In the following, we look at the results from various different perspectives.

Figure 2.4 shows the estimated impulse response functions for the variables in the benchmark specification for three selected quarters at the beginning (1980Q4), towards the middle (1995Q4), and at the end of the sample (2008Q4). The results show that the estimated short-run multiplier on output is larger at the beginning of the sample, the point estimate being around 0.7 for 1980Q4 compared to 0.4 for 2008Q4. Moreover,

---

14An output multiplier smaller than one combined with (marginally) positive point estimates for consumption and investment could be explained by a decline in net exports, although we have not included this variable as it is not available at the euro area level. This explanation is however consistent with SVAR results for a panel of 14 EU countries discussed in Beetsma, Giuliodori, and Klaassen (2008), showing that, on average, the trade balance falls by 0.5% of GDP on impact due to a 1% of GDP increase in government spending.
the effects of spending shocks on output seem to have lost persistence over time, and they are increasingly negative at longer horizons. In particular, the estimated response of GDP at a horizon of five years is about -0.7 percent for 1980Q4 but -1.6 percent for 2008Q4. The time-varying techniques thus indicate increasingly contractionary longer-term effects of a spending expansion. Furthermore, while the initial output response is positive with 68 percent probability in the initial period, the probability bands include the zero line at the end of the sample. Instead, the response after five years is significantly negative only in the most recent period. The results further suggest that the effects on consumption have decreased over time in a similar way as the effects on
output. We also note a stronger response of the nominal interest rate.

The conclusions from Figure 2.4 are confirmed in Figure 2.5, which shows the estimated state-dependent median impulse responses for each year in the sample. Only the fourth-quarter response in each year is reported, such that the first impulse response refers to 1980Q4 while the last one refers to 2008Q4. The figure shows that the estimated short-run effects of spending shocks on output and consumption are largest towards the end of the 1980s and lowest towards the recent period, whereas the estimated effects at longer horizons are steadily falling from the beginning towards the end of the sample. The estimated impulse response of government spending itself is however rather stable over time.
Figure 2.6: Time-varying impulse responses III – selected horizons

Notes. See Figure 2.4; $k$-th horizon median responses for $k = 0, 4, 20$.

Figure 2.6 shows the responses of all variables over time at selected horizons, i.e. the contemporaneous responses, the responses after one year, and the responses after five years. The estimated contemporaneous multiplier on output is slightly below one for the period 1980-1985, larger than one for the period 1986-1990, and then falls over the period 1991-2003 to reach values around 0.5 in 2004-2008. At the five-year horizon, the estimated effects on output and consumption of an initial 1% of GDP expansion are substantially lower for the recent decade, from -1.4% to -1.7% of GDP, compared to -0.7% to -1% in the 1980s. In general, the changes in the effects on output are similar as the changes in the effects on private consumption. The estimated contemporaneous reaction of the nominal interest rate is negative from 1980 until around 1999-2002.
Figure 2.7: Pair-wise joint posterior distributions of time-varying impulse responses

Notes. See Figure 2.4; pair-wise responses of output, consumption (in percent of GDP), and the interest rate (in percentage points), computed across their posterior distribution at horizons $k = 0, 4, 20.$

and positive afterwards. The estimated interest rate response at longer horizons also increases over time. A stronger interest rate response thus might have contributed to the decrease in fiscal multipliers.

To test differences in the above responses over time, we compute the joint pair-wise distributions of impulse responses at two selected horizons. That is, in Figure 2.7 (sorted) draws from the posterior distribution of output and consumption responses and the interest rate response in 1980Q4 are plotted against draws for 2008Q4.$^{15}$ Results are

$^{15}$A similar exercise is implemented in Cogley et al. (2010). There are many alternative pairs of quarters to choose from, but the results are not particularly sensitive to this choice as long as the
Figure 2.8: Time-varying impulse responses IV – extended specifications

![Graphs showing time-varying impulse responses for Investment, Net Taxes, Real Wage, and Inflation.](image)

**Notes.** See Figure 2.4; only median impulse responses are reported.

reported for the impact responses, the one-year responses, and the five-year responses. Each point in the respective panels represents a draw from the joint distribution for 1980Q4 and 2008Q4. Thus, combinations near the 45 degree line represent pairs for which there was little or no change over time and those above (below) the 45 degree line are pairs where the response of the respective variable has increased (decreased). The figure shows that the lower tails of the distributions of the output and consumption responses have shifted downwards, especially at longer horizons, whereas the upper tails appear comparably stable. Therefore, the median estimates have shifted downwards as well. Regarding the interest rate, both time variation in its impact response and the response after five years turn out to be important.

periods used are sufficiently distant from each other.
We also investigate time variation in the effects of spending shocks on a broader set of macroeconomic indicators, adding one at a time private investment, net taxes, the real wage, and the HICP-based inflation rate to the estimated VAR. Figure 2.8 shows the estimated state-dependent median impulse responses. We observe a small positive short-term effect on private investment and a crowding out at longer horizons. Similarly as the multiplier on output and the effect on consumption, the effect on investment was larger in the first part of the sample. The reaction of net taxes to government spending shocks has remained comparably stable over time, and throughout the response is smaller than 1% of GDP, indicating that the primary deficit has always increased due to the spending shock. A smaller overall fiscal expansion can thus not serve as an explanation for smaller spending multipliers.

The response of the real wage is estimated to be positive for several quarters after the shock throughout the sample, but it shows a larger initial reaction and a more persistent response in the first part of the sample towards the late 1980s. The initial response of inflation was close to zero throughout, but we observe a stronger medium-term response during the 1980s and most of the 1990s. As the nominal interest rate reacts more strongly to government spending shocks, this result implies that the real interest rate response has tended to increase over time.

2.5 Robustness checks

This section reports the results of several robustness checks, as listed below.

2.5.1 Scaling factors

In the estimation of the TVP-VAR, we have elicited relative conservative priors on time variation, in particular the scaling factors \( k_Q, k_S, \) and \( k_W \) which parameterize the priors on the covariance matrices of the shocks in the state equations, as described in Appendix 2.B. The values were \( k_Q = k_W = 10^{-4} \) and \( k_S = 10^{-2} \), following the related literature (see, in particular, Primiceri, 2005). To check the sensitivity of the estimation outcomes, we now further reduce the scaling factors one at a time to \( k_Q = k_W = 0.5 \times 10^{-4} \), and \( k_S = 0.5 \times 10^{-2} \), keeping the other two factors fixed at
their benchmark values in each case.

The results are summarized in Figures 2.9, 2.10, and 2.11, respectively. Figure 2.9 shows that the reduction of the coefficients scaling factor \( k_Q \) especially increases the estimated time variation of the short-run multiplier on output and the contemporaneous effect on consumption. It therefore seems that, compared to the previous results, some of the time variation in the VAR coefficients is instead picked up by the covariance terms. On the other hand, Figures 2.10 and 2.11 show that the reductions of \( k_S \) and \( k_W \) only lead to relatively small changes in the amount and the direction of the estimated time variation.
2.5.2 Stationarity conditions

In their analysis of U.S. monetary policy, Cogley and Sargent (2002) have proposed to discard draws from the Gibbs sampler that do not satisfy stationarity conditions, and many related studies have followed this approach. However, we have argued above that the stationarity restriction is harder to defend for aggregate euro area fiscal data since fiscal variables may have followed unstable paths in some countries. The potential downside of not imposing the stationarity conditions is that this may exaggerate the amount of time variation due to a potentially large amount of unstable draws. We therefore check the robustness of the TVP-VAR results when stationarity conditions are
imposed on the VAR coefficients. Formally, the random walk process 2.3 for the VAR coefficients $\beta_t, t = 1, 2, 3, \ldots, T$, characterizes the conditional density $f(\beta_t|\beta_{t-1}, Q)$. Following Cogley and Sargent (2002), introduce an indicator function $I(\beta_t)$ which rejects unstable draws that do not satisfy standard eigenvalue stability conditions and which thus enforces stationarity of the estimated TVP-VAR at each point of time. The VAR coefficients are thus postulated to evolve according to

$$p(\beta_t|\beta_{t-1}, Q) = I(\beta_t)f(\beta_t|\beta_{t-1}, Q).$$

Figure 2.12 shows the estimated state-dependent effects at selected horizons when
Figure 2.12: Robustness IV – stationarity conditions imposed

![Figure 2.12: Robustness IV – stationarity conditions imposed](image)

Note. See Figure 2.6.

the stationarity conditions are imposed. A comparison with the previous results indicates no significant differences to the benchmark case. The multipliers show somewhat less high-frequency variation but the broad patterns are similar.

2.5.3 Anticipation effects

To check for the possible presence of anticipation effects, this section confronts the estimated SVAR shocks with macroeconomic forecasts to see whether the identified shocks are potentially predictable. This exercise follows Ramey (2011b) who shows that, for the U.S., SVAR spending shocks are Granger-caused by forecasts made one
to four quarters earlier (i.e. they are predictable). Thus, we perform tests of Granger causality from various variables conveying information about future policy and macro developments onto the time series of estimated SVAR spending shocks. We use both survey data from Consensus Economics, as in Ramey (2011b), and publicly available short-term forecasts by the European Commission. The Consensus data summarizes the predictions of professional forecasters at banks and other financial institutions. This data is thus taken to represent economic agents’ (or market participants’) expectations on future macroeconomic developments. The European Commission forecasts do not directly reflect such expectations, but they do cover a longer period than the survey data, thus increasing the power of the tests.\textsuperscript{16} We therefore exploit both data sets.

The exercise conducted below is however subject to the following limitations. First, we are forced to use time-aggregated quarterly data in the estimation since the macroeconomic forecasts are only available on an annual basis. Second, we also need to restrict the analysis to the fixed parameters VAR as the number of observations in the time-aggregated data is not sufficient to estimate the TVP-VAR. Third, the data incorporates predictions on government deficits and deficit-to-GDP ratios as the only fiscal variables instead of direct forecasts on government spending. The results reported below should be interpreted with these limitations in mind.

The results of the Granger tests are reported in Table 2.1. Following Ramey (2011b), the SVAR shocks in period $t$ are regressed on a constant, their own lags and various forecasts made in period $t-1$ for period $t$.\textsuperscript{17} The null hypothesis is that the forecasts do not Granger-cause the SVAR shocks.\textsuperscript{18} The first panel of Table 2.1 shows that the null hypothesis cannot be rejected at the 10 percent significance level for any of the European Commission’s forecasts in isolation, on the deficit-to-GDP ratio and real GDP growth, and also not if both forecasts are included as right-hand side variables. Similarly, the second panel shows that the null hypothesis cannot be

\textsuperscript{16}The European Commission provides forecasts in November of every year for the following year since the 1970s for a number of European countries. Consensus Economics provides forecasts every month for 1991 onwards. Forecasts on the budget deficit are only available for 1994 onwards.

\textsuperscript{17}The results are robust to the use of additional lagged values of the left-hand side and/or the right-hand side variables, as well as the addition of the period $t$ variables (and their lagged values) which are included in the VAR model on the right-hand side.

\textsuperscript{18}The Granger causality test is identical to an F-test of the null hypothesis that the unrestricted model, which includes the forecasts, does not provide a better fit than the restricted model, which excludes the forecasts.
Table 2.1: Granger causality tests using macroeconomic forecasts\textsuperscript{a}

<table>
<thead>
<tr>
<th>Hypothesis test\textsuperscript{b}</th>
<th>F-statistic</th>
<th>10% critical value</th>
<th>Conclusion (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>European Commission Forecasts\textsuperscript{c}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deficit-to-GDP ratio forecasts → SVAR shocks</td>
<td>0.004</td>
<td>2.949</td>
<td>No (0.949)</td>
</tr>
<tr>
<td>GDP growth forecasts → SVAR shocks</td>
<td>0.001</td>
<td>2.949</td>
<td>No (0.973)</td>
</tr>
<tr>
<td>All forecasts → SVAR shocks</td>
<td>0.002</td>
<td>2.575</td>
<td>No (0.998)</td>
</tr>
<tr>
<td>Deficit-to-GDP ratio forecasts → actual spending growth</td>
<td>4.894</td>
<td>2.949</td>
<td>Yes (0.038)</td>
</tr>
<tr>
<td>Consensus Economics Forecasts\textsuperscript{c}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deficit growth forecasts → SVAR shocks</td>
<td>0.027</td>
<td>3.225</td>
<td>No (0.872)</td>
</tr>
<tr>
<td>GDP growth forecasts → SVAR shocks</td>
<td>0.373</td>
<td>3.102</td>
<td>No (0.551)</td>
</tr>
<tr>
<td>Consumption growth forecasts → SVAR shocks</td>
<td>0.155</td>
<td>3.102</td>
<td>No (0.700)</td>
</tr>
<tr>
<td>Interest rate forecasts → SVAR shocks</td>
<td>0.785</td>
<td>3.102</td>
<td>No (0.391)</td>
</tr>
<tr>
<td>All forecasts → SVAR shocks</td>
<td>0.049</td>
<td>2.693</td>
<td>No (0.995)</td>
</tr>
<tr>
<td>Deficit growth forecasts → actual spending growth</td>
<td>0.320</td>
<td>3.225</td>
<td>No (0.320)</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The first variable at time \( t \) is regressed on a constant, its own lag at time \( t - 1 \), and the forecast made at time \( t - 1 \) of the second variable for period \( t \).
\textsuperscript{b} The null hypothesis is that the second variable does not Granger-cause the first variable.
\textsuperscript{c} For the European Commission forecasts (1982-2006), GDP is measured as real annual growth rate and the deficit-to-GDP ratio is measured in nominal terms. For the Consensus Economics forecasts (1992-2008), all variables except the interest rate are measured as real annual growth rates, using consumer prices as deflators, and the interest rate is measured in nominal terms. See Appendix 2.D for details on the data definitions.
rejected for the professional forecasts on the growth rates of the budget deficit, GDP, private consumption, and the short-term interest rate. In addition, we check whether the Commission’s forecasts on the deficit-to-GDP ratio and professional forecasts on the budget deficit Granger-cause realized spending growth. This is the case for the Commission’s forecasts, where the null hypothesis is rejected. Thus, although the forecasts do have predictive power for realized spending, they do not predict the SVAR spending shocks. Overall, this exercise does not provide strong reasons to doubt the validity of the identification approach due to anticipation effects.

2.6 The fiscal transmission mechanism

This section exploits the results obtained in the previous step with the aim of providing empirical evidence on the determinants of the effects of government spending shocks in the euro area. Section 2.6.1 reviews the main theories on the fiscal transmission mechanism, focusing on (i) the level of government debt, (ii) asset market participation and the availability of credit, (iii) the degree of trade openness, (iv) the share of government investment in total spending, and (v) the wage component of total spending. Section 2.6.2 relates these factors to the estimated effects of spending shocks using regression analysis.

2.6.1 Views on the transmission mechanism

(i) Government debt. Experience from past fiscal consolidations suggests the possibility that in times of fiscal stress an economy’s response to fiscal shocks changes. That is, positive consumption growth was observed after prolonged and substantial deficit cuts. This is the hypothesis of “expansionary fiscal contractions” brought about by Giavazzi and Pagano (1990).\(^\text{19}\) Indeed, for a panel of 19 OECD countries, Perotti (1999) finds that the effect of spending shocks on consumption can be positive if the

\(^{19}\text{See also Giavazzi, Jappelli, and Pagano (2000). Giavazzi and Pagano (1990) study episodes of large fiscal consolidations in Denmark during 1983-1986 and in Ireland during 1987-1989. In these episodes the cyclically adjusted deficit as a share of GDP declined by 9.5 percent and 7.2 percent relative to the preconsolidation year and yet private consumption increased by 17.7 percent and 14.5 percent cumulatively. Alesina and Perotti (1996) identify similar episodes in several other European countries and Canada during the 1980s.}\)
initial financing needs of the government are small, arguing that this outcome is due to the convexity of tax distortions: a (larger) expected increase in taxation tomorrow causes a (larger) decline in wealth and a (larger) fall in consumption today.

(ii) Credit. In standard general equilibrium models, expansionary government spending shocks tend to generate a crowding out of private consumption. The reason is the negative wealth effect induced by higher future tax payments, which increases consumer saving due to the consumption smoothing objective. However, credit constraints and limited asset market participation may dampen this effect. For example, Galí et al. (2007) show that a spending shock can generate an increase in aggregate consumption in a New Keynesian model conditional on, in particular, a relatively large fraction of liquidity-constrained consumers. In addition, it has recently been argued that fiscal stabilization policy may be more effective during recessions since credit constraints might then bind across a wider range of agents. In particular, Roeger and 't Veld (2009) allow for credit-constrained households along the lines of the financial accelerator literature, thus allowing the stringency of credit constraints to vary over the cycle, and show that stabilization policy becomes more effective since the propensity to consume out of current income increases during recessions.20

(iii) Openness. It is often claimed that the effectiveness of fiscal policy depends on the degree of openness to trade.21 The argument is that in very open economies domestic output will be comparatively less affected by a fiscal expansion since a large fraction of the intended stimulus falls on imports. For instance, Beetsma, Giuliodori, and Klaassen (2008) show that a 1% of GDP increase in public spending in the EU leads to a fall of the trade balance by 0.5% of GDP on impact and a peak fall of 0.8% of GDP. With respect to time variation in fiscal multipliers, the effects of an increase in spending on GDP are then expected to be smaller the higher the degree of openness. Below we use the import share as a proxy for the degree of openness since imports

20Tagkalakis (2008) also provides evidence for asymmetric effects of fiscal policy for a panel of 19 OECD countries over the period 1970-2002, showing that a spending shock has a larger effect on private consumption in downturns than in upturns. See also Auerbach and Gorodnichenko (2010).
21See, for instance, Perotti (2005) who however argues that the increase in openness is probably too small to account for the decline in spending multipliers in OECD economies.
are the relevant channel through which openness to trade may affect fiscal multipliers according to this argument.

(iv) Government investment. Although not all empirical studies find a growth-enhancing effect of public capital, there is now more consensus than in the past that public capital supports economic growth (see Romp and de Haan, 2007). A corresponding change in the composition of spending may therefore contribute to changing spending multipliers. Macroeconomic models which account for productive public capital typically predict that increases in government investment can generate larger fiscal multipliers than increases in government consumption, due to the beneficial aggregate supply effect of public capital. 22 On the other hand, Leeper, Walker, and Yang (2010) have recently provided evidence showing that government investment projects in the U.S. are subject to substantial implementation lags. Private investment and employment are then postponed until the public capital is on line, which, as these authors show in a macroeconomic model, can lead to smaller short-term multipliers.

(v) Wage component. More than half of government consumption in the euro area consists of wage payments to government employees, whereas less than half consists of goods purchases. Several studies emphasize that this distinction matters when assessing the impact of spending shocks on the macroeconomy. For example, based on a neoclassical model, Finn (1998) shows that government employment shocks raise the real wage and thus act as a transfer to households, which dampens the wealth effects on consumption and labor supply. 23 Using SVAR analysis, Perotti (2008) shows that, in the U.S., government employment shocks have larger effects on output and consumption than shocks to government goods purchases. On the other hand, Alesina and Ardagna (2010) argue that in an imperfect labor market a decrease in government employment could reduce job finding probabilities, whereas a decrease in government wages could decrease incomes of workers in the public sector. In both cases, reservation utilities and wages demanded for private sector workers would decrease, which

---

22 See, for instance, Baxter and King (1993), Pappa (2005), and Straub and Tchakarov (2007).
23 Pappa (2005) demonstrates that government employment shocks have similar effects in a New Keynesian framework.
may increase profits, investment, and competitiveness.

### 2.6.2 Driving forces of time variation

Several testable hypotheses can be derived from the discussion in Section 2.6.1. First, the effects of spending shocks on output and consumption are expected to be smaller the higher the initial debt-to-GDP ratio. Second, the effects can be higher if households are more restricted in their access to credit, or if actual output is below potential output. Third, a higher share of imports over GDP is expected to lead to smaller spending multipliers. Fourth, a higher government investment share can lead to higher spending multipliers, but if implementation lags play a role, short-term multipliers can also be smaller. Fifth, a higher wage share can result in larger or smaller effects on economic activity depending on the degree of labor market competitiveness.

The above hypotheses are analyzed using regression inference. We apply Bayesian linear regressions, using the estimated time-varying effects on output and consumption as dependent variables. This type of two-step approach, while based here on time-varying parameters, is close in spirit to Fatás and Mihov (2006). We distinguish both contemporaneous effects and longer-term effects after five years. Further, since the dependent variables are themselves estimated parameters, the standard errors of the regression coefficients are adjusted to account for the uncertainty in the dependent variables. Not doing so may give a biased view on the importance of the restrictions implied by the explanatory variables and might thus artificially produce significant effects even when the “true” ones are negligible (see Canova and Pappa, 2006). In particular, we use each of 1,000 posterior draws of multipliers from the identified TVP-VAR model in turn as dependent variable. We then generate 1,100 draws of regression coefficients by Gibbs sampling and omit the first 100 draws for each regression. This

---

24 Diffuse normal priors with mean zero and standard deviation $10^6$ are specified for the regression coefficients. All regressions include a constant and a linear trend to address possible concerns of spurious causation. Using a linear-quadratic trend instead of a linear trend did not lead to significant changes in the results. We also account for the possible presence of heteroskedastic disturbances, where we use diffuse priors on the variance terms. The regressions are estimated using a Gibbs sampling algorithm with 1,100 draws, dropping the first 100 draws, see Geweke (1993).

25 Fatás and Mihov (2006) study the determinants of output elasticities of government spending. The latter are estimated in a first step over a sample of 48 U.S. states. In a second step, the authors analyze the impact of different fiscal rules on those elasticities.
Figure 2.13: Potential determinants of spending multipliers

Notes. The debt-to-GDP ratio is in nominal annual terms; the ratio of credit to households over GDP is outstanding (end-of-period) loans to households divided by the sum of nominal GDP of the last four consecutive quarters; the output gap is measured as quarterly percentage deviation from trend real GDP, trend is based on HP filter with smoothing parameter 1600; the ratio of imports over GDP and the shares of government investment and wage expenditures in total spending are based on quarterly nominal data; source of fiscal data: Paredes et al. (2009); source of remaining data: ECB’s Area-Wide Model database and Bank of International Settlements macroeconomic series (data on loans).

leaves us with 1,000,000 posterior draws from the posterior distribution of regression coefficients, conditional on the full posterior distribution of estimated multipliers, from which we compute means and posterior probabilities.

Figure 2.13 shows the explanatory variables used in the regression analysis. The lagged aggregate euro area debt-to-GDP ratio is used to measure the initial financing needs of euro area governments. Availability of credit is measured by the lagged ratio of credit to households over GDP. The state of the business cycle is approximated by the lagged HP-filtered output gap. Lagged values are used to address potential reverse causation from spending multipliers on output and the business cycle. The ratio of
imports over GDP (in lagged terms) is used to assess the impact of changes in the degree of openness. Finally, we include the contemporaneous shares of government investment and employee compensation over total spending to assess the impact of changes in the composition of spending on its overall effects.

The results for contemporaneous effects and the effects after five years, respectively, are reported in Tables 2.2 and 2.3. The point estimates of the regression coefficients are the means of their posterior distribution. The statistical significance of the regression coefficients is measured as the posterior probability that they are non-positive (non-negative) if their point estimates are positive (negative).

The results in Table 2.2 show that, on average, an increase in the debt-to-GDP ratio has a negative but small effect on short-term multipliers. On the other hand, a rise in the credit ratio is estimated to have a larger impact, a one percentage point increase leading on average to a decline in the spending effect on output (consumption) between 0.04 and 0.06 points (between 0.02 and 0.04 points). The credit ratio has increased from 30 percent in 1980 to almost 60 percent in 2008, such that increasing credit availability is estimated to have contributed substantially to the decline in contemporaneous multipliers. The output gap enters with an unexpected positive sign, whereas a rise in the import share is estimated to have a negative but mostly insignificant effect. The estimated impact of an increase in the share of government investment in total spending is positive whereas an increase in the share of wage payments is estimated to have a negative effect. In the largest regression model for output (consumption), a unitary increase of the investment share is estimated to cause an average increase in the contemporaneous effects by 0.07 points (0.04 points). A unitary increase in the wage share leads to a decrease in the effects by 0.04 points (0.03 points).

The evidence presented in Table 2.3, on the other hand, suggests that the level of government debt relative to GDP is the main determinant of the longer-term effects of government spending. For both output and consumption, a one percentage point increase in the debt ratio leads on average to a decline by 0.01 points in the associated effects, the coefficients being negative with at least 95 percent probability in all regression models. The coefficients on some of the remaining variables do have the expected signs, but none of them are different from zero with more than 90 percent probability.
Table 2.2: Bayesian linear regressions on contemporaneous effects\textsuperscript{a,b}

<table>
<thead>
<tr>
<th></th>
<th>Multiplier on output</th>
<th></th>
<th>Effect on consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Gov. debt/GDP (-1)</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Credit/GDP (-1)</td>
<td>-0.06***</td>
<td>-0.06***</td>
<td>-0.04***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Output gap (-1)</td>
<td>0.03**</td>
<td>0.05***</td>
<td>0.04**</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Imports/GDP (-1)</td>
<td>-0.02*</td>
<td>-0.01*</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Inv. share</td>
<td>0.03*</td>
<td>0.07**</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Wage share</td>
<td></td>
<td>-0.04**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.71</td>
<td>3.11***</td>
<td>2.99***</td>
</tr>
<tr>
<td></td>
<td>(1.37)</td>
<td>(1.10)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>Trend</td>
<td>-0.01***</td>
<td>0.01***</td>
<td>0.01***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Observations</td>
<td>112</td>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The Bayesian regressions allow for heteroskedastic errors following Geweke (1993). The standard error adjustment proceeds by using each of 1,000 multipliers in the posterior distribution from the identified TVP-VAR as dependent variable. All regressions are then estimated using a Gibbs sampling algorithm with 1,100 draws and 100 omitted draws. This leaves us with 1,000,000 posterior draws of regression coefficients.

\textsuperscript{b} The point estimates are the posterior means of the posterior distribution. Standard deviations are reported in parentheses. Asterisks indicate posterior probabilities that the regression coefficients are non-positive if the point estimates are positive or non-negative if the point estimates are negative (*less than ten percent, **less than five percent, ***less than one percent). Explanatory variables are measured in percent.
Table 2.3: Bayesian linear regressions on effects after five years\textsuperscript{a,b}

<table>
<thead>
<tr>
<th>Multiplier on output</th>
<th>Effect on consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gov. debt/GDP (-1)</td>
<td></td>
</tr>
<tr>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>-0.01**</td>
<td>-0.01**</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>-0.01**</td>
<td>-0.01**</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>-0.01**</td>
<td>-0.01**</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>-0.01**</td>
<td>-0.01**</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Credit/GDP (-1)</td>
<td></td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.09)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.09)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>-0.03</td>
<td>-0.02</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>-0.03</td>
<td>-0.02</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>(0.11)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Output gap (-1)</td>
<td></td>
</tr>
<tr>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>(0.11)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Output gap (-1)</td>
<td></td>
</tr>
<tr>
<td>Imports/GDP (-1)</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Inv. share</td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(0.13)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Wage share</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
</tr>
<tr>
<td>-0.67</td>
<td>-0.56</td>
</tr>
<tr>
<td>(1.96)</td>
<td>(1.69)</td>
</tr>
<tr>
<td>0.02</td>
<td>-0.10</td>
</tr>
<tr>
<td>(3.24)</td>
<td>(1.76)</td>
</tr>
<tr>
<td>0.03</td>
<td>-0.09</td>
</tr>
<tr>
<td>(2.86)</td>
<td>(1.60)</td>
</tr>
<tr>
<td>0.05</td>
<td>-0.07</td>
</tr>
<tr>
<td>(2.87)</td>
<td>(1.54)</td>
</tr>
<tr>
<td>0.08</td>
<td>-0.10</td>
</tr>
<tr>
<td>(2.92)</td>
<td>(3.70)</td>
</tr>
<tr>
<td>-0.36</td>
<td>-0.82</td>
</tr>
<tr>
<td>(4.50)</td>
<td>(7.34)</td>
</tr>
<tr>
<td>Trend</td>
<td></td>
</tr>
<tr>
<td>-0.01***</td>
<td>-0.01***</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>-0.01***</td>
<td>-0.01***</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>-0.01***</td>
<td>-0.01***</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>-0.01***</td>
<td>-0.01***</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>112</td>
</tr>
<tr>
<td>112</td>
<td>112</td>
</tr>
<tr>
<td>112</td>
<td>112</td>
</tr>
<tr>
<td>112</td>
<td>112</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The Bayesian regressions allow for heteroskedastic errors following Geweke (1993). The standard error adjustment proceeds by using each of 1,000 multipliers in the posterior distribution from the identified TVP-VAR as dependent variable. All regressions are then estimated using a Gibbs sampling algorithm with 1,100 draws and 100 omitted draws. This leaves us with 1,000,000 posterior draws of regression coefficients.

\textsuperscript{b} The point estimates are the posterior means of the posterior distribution. Standard deviations are reported in parentheses. Asterisks indicate posterior probabilities that the regression coefficients are non-positive if the point estimates are positive or non-negative if the point estimates are negative (*less than ten percent, **less than five percent, ***less than one percent). Explanatory variables are measured in percent.
To summarize, the second-stage regressions indicate that (i) a higher level of government debt relative to GDP is associated with lower spending multipliers at longer horizons. (ii) The ratio of credit over GDP seems to be an important determinant of the observed time variation in the short-run effects of spending shocks. (iii) The degree of openness, measured here by the share of imports over GDP, does not seem to be an important driving force of spending multipliers. With respect to compositional effects, (iv) a higher share of government investment in total spending has a positive effect on the size of short-run multipliers, whereas (v) a larger wage component of government spending is associated with smaller short-run multipliers.

2.7 Conclusion

This chapter has estimated vector autoregressions with drifting coefficients and stochastic volatility for the euro area, with the aim of investigating changes in the effects of government spending shocks over the period 1980-2008 and, based on second-stage inference, revealing the driving forces of the fiscal transmission mechanism.

Our results indicate that the effectiveness of spending shocks in stimulating economic activity has decreased over time. The estimated short-run multipliers are highest in the late 1980s when they reached values above unity, but they fall afterwards to values closer to 0.5 in the current decade. Longer-term multipliers show a more than two-fold decline since the 1980s. These results suggest that other components of aggregate demand are increasingly being crowded out by spending-based fiscal expansions. In particular, the response of private consumption to government spending shocks has become substantially weaker over time. We also document a weaker response of real wages, whereas the nominal interest rate shows a stronger reaction.

With respect to the driving forces of time variation, our evidence points towards availability of credit as one of the main determinants of the short-run effects of government spending. Furthermore, a lower share of government investment and a larger wage component in total spending seem to have contributed to the observed decline in short-run multipliers. Finally, our results suggest that rising government debt is associated with declining spending multipliers at longer horizons, and thus increasingly
negative longer-term consequences of fiscal expansions.

2. A Details of the Gibbs sampler

This appendix outlines the details of the Gibbs sampling algorithm used for estimation of the TVP-VAR model. The algorithm generates a Markov chain which is a sample from the joint posterior distribution of the VAR parameters (i.e. coefficient states, covariance states, volatility states, and hyperparameters). It combines elements of Benati and Muntaz (2007), Cogley and Sargent (2005), and Primiceri (2005), with a few additional restrictions on the structure of the hyperparameters.

In the following, $x^t$ denotes the history of $x$ up to time $t$, i.e. $x^t = [x_1', x_2', x_3', \ldots, x_t']'$, and $T$ denotes the sample length. Furthermore, re-write the observation equation (2.2) in the main text conveniently as

$$y_t = X_t' \beta_{t-1} + u_t, \quad (2.6)$$

where $X_t' = I \otimes [y_{t-1}', y_{t-2}', y_{t-3}', \ldots, y_{t-p}', z_t']$. The estimation of the model proceeds in the following four steps.

(i) **Drawing coefficient states** $\beta^T$. Conditional on $A^T$ and $H^T$, one obtains a history $R^T$. Then, conditional on $y^T$, $R^T$, and $Q$, the observation equation (2) is linear with Gaussian innovations and a known covariance matrix. The posterior density of the coefficients can be factored as

$$f(\beta^T|y^T, R^T, Q) = f(\beta_T|y^T, R^T, Q) \prod_{t=1}^{T-1} f(\beta_t|\beta_{t+1}, y^t, R_t, Q), \quad (2.7)$$

where

$$\beta_t|\beta_{t+1}, y^t, R^T, Q \sim N(\beta_{t|t+1}, P_{t|t+1}),$$

$$\beta_{t|t+1} = E[\beta_t|\beta_{t+1}, y^t, R^T, Q],$$

$$P_{t|t+1} = E[P_t|P_{t+1}, y^t, R^T, Q].$$

---

26 Conditioning factors which are redundant in the respective step are omitted.
The conditional means and variances can be computed using the Kalman filter and a backward recursion (see Carter and Kohn, 1994). The Kalman filter delivers

\[ P_{t|t-1} = P_{t-1|t-1} + Q, \quad K_t = P_{t|t-1}X_t'(X_tP_{t-1|t-1}X_t + R_t)^{-1}, \]
\[ \beta_{t|t} = \beta_{t-1|t-1} + K_t(y_t - X_t'\beta_{t-1|t-1}), \quad P_{t|t} = P_{t-1|t-1} - K_tX_t'P_{t|t-1}. \]

The initial values \( \beta_{0|0} \) for this recursion are the OLS point estimates from the initial sample, and the initial value \( P_{0|0} \) is their covariance matrix. The initial \( R_t \) is the OLS covariance matrix of the reduced-form VAR model. The covariance matrix \( Q \) is a scaled version of the variance-covariance matrix of the coefficients. The Kalman filter delivers as its last points \( \beta_{T|T} \) and \( P_{T|T} \). Draws from (2.7) are then obtained by a backward recursion. The first point in the backward recursion is a draw from \( N(\beta_{T|T}, P_{T|T}) \). The remaining draws are from \( N(\beta_{t|t+1}, P_{t|t+1}) \), where the means and variances are derived as follows:

\[ \beta_{t|t+1} = \beta_{t|t} + P_{t|t}P_{t+1|t}^{-1}(\beta_{t+1} - \beta_{t|t}), \quad P_{t|t+1} = P_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t}. \]

(ii) **Drawing covariance states** \( A^T \). Conditional on \( y^T, \beta^T, \) and \( H^T \), the system of equations (2.6) can be written as follows:

\[ A_t(y_t - X_t'\beta_t) = A_t\hat{y}_t = H_t^{1/2}v_t. \]  

Moreover, \( A_t \) is lower diagonal (with ones on the main diagonal) such that (2.8) can be re-written as

\[ \hat{y}_t = Z_t\alpha_t + H_t^{1/2}v_t, \]  

where \( \alpha_t \) is defined as in the main text and \( Z_t \) has the structure

\[
Z_t = \begin{bmatrix}
0 & \cdots & \cdots & 0 \\
-\hat{y}_{1,t} & 0 & \cdots & \vdots \\
0 & (-\hat{y}_{1,t}, -\hat{y}_{2,t}) & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & (-\hat{y}_{1,t}, \cdots, -\hat{y}_{n-1,t})
\end{bmatrix},
\]
where \( n \) denotes the number of variables in the VAR model. The system of equations (2.9) has a Gaussian but non-linear state-space form. However, under the assumption of (block) diagonality of \( S \) the problem becomes linear (see Primiceri, 2005). The forward (Kalman filter) and backward recursions of the previous step can then be applied equation by equation. Hence, the procedure allows to recover \( \alpha^T \) through

\[
\alpha_{i,t+1} = E[\alpha_{i,t} | \alpha_{i,t+1}, y^t, \beta^T, H^T, S_i],
\]

\[
\Lambda_{i,t+1} = \text{var}[\alpha_{i,t} | \alpha_{i,t+1}, y^t, \beta^T, H^T, S_i],
\]

where \( \alpha_{i,t} \) is the block of \( \alpha_t \) corresponding to the \( i \)-th equation and \( S_i \) is the associated \( i \)-th block of \( S \). The initial values for the Kalman filter are obtained from a decomposition of the OLS covariance matrix.

(iii) Drawing volatility states \( H^T \). To sample the stochastic volatilities, the univariate algorithm of Jacquier, Polson, and Rossi (1994) is applied to each element of \( H_t \). The orthogonalized residuals \( v_t = A_t u_t \) are observable conditional on \( y^T, \beta^T, \) and \( A^T \). We can use the univariate setting since the stochastic volatilities are assumed to be independent, following Cogley and Sargent (2005). Jacquier, Polson, and Rossi (1994) show that the conditional kernel is

\[
f(h_{i,t}|h_{i,t-1}, h_{i,t+1}, v^T_i, w_i) \propto f(h_{i,t}|h_{i,t-1}, h_{i,t+1}, v^T_i, w_i),
\]

where \( w_i \) is the \( i \)-th diagonal element of \( W \) and \( h_{-i,t} \) represents the vector of \( h \)'s at all other dates. Using Bayes’ theorem, the conditional kernel can be expressed as

\[
f(h_{i,t}|h_{i,t-1}, h_{i,t+1}, v^T_i, w_i) \propto f(u_{i,t}|h_{i,t})f(h_{i,t}|h_{i,t-1})f(h_{i,t+1}|h_{i,t})
\]

\[
\propto h_{i,t}^{-1.5} \exp \left( -\frac{v_{i,t}^2}{2h_{i,t}} \right) \exp \left( -\frac{(\ln h_{i,t} - \mu_{i,t})^2}{2\sigma_{ic}^2} \right),
\]

where \( \mu_{i,t} \) and \( \sigma_{ic}^2 \) are the conditional mean and the conditional variance of \( h_{i,t} \) implied by equation (2.5) in the main text and knowledge of \( h_{i,t-1} \) and \( h_{i,t+1} \). For a geometric
random walk these parameters are

$$\mu_{i,t} = 0.5(\log h_{i,t-1} + \log h_{i,t+1}) \quad \text{and} \quad \sigma_{ic}^2 = 0.5w_i.$$  

In practice $h_{i,t+1}$ is taken from the previous Gibbs iteration.\(^{27}\) Jacquier, Polson, and Rossi (1994) propose a Metropolis step instead of a Gibbs step, because the normalizing constant is expensive to calculate in (2.10). Hence, one draws from a stand-in density and then uses the conditional likelihood $f(u_{i,t}|h_{i,t})$ to calculate the acceptance probability for that draw. Cogley and Sargent (2005) suggest to use the log-normal density implied by equation (2.5) in the main text as the stand-in density:

$$g(h_{i,t}) \propto h_{i,t}^{-1} \exp \left( -\frac{(\log h_{i,t} - \mu_{i,t})^2}{2\sigma_{ic}^2} \right).$$

The acceptance probability for the $m$-th draw is

$$q_m = \frac{f(v_{i,t}|h_{i,t}^m)g(h_{i,t}^m)}{g(h_{i,t}^{m-1})} \frac{g(h_{i,t}^{m-1})}{f(v_{i,t}|h_{i,t}^{m-1})g(h_{i,t}^{m-1})} = \frac{(h_{i,t}^m)^{-1/2} \exp \left( -\frac{v_{i,t}^2}{2h_{i,t}^m} \right)}{(h_{i,t}^{m-1})^{-1/2} \exp \left( -\frac{v_{i,t}^2}{2h_{i,t}^{m-1}} \right)},$$

where $h_{i,t}^m = h_{i,t}^{m-1}$ if the draw is rejected. This algorithm is applied on a date-by-date basis to each element of $u_t$. The formulas are slightly different for the first and last element. For the first element we have

$$\mu_{i1} = \sigma_{ic}^2 \left( \frac{\mu_{i0}}{\sigma_{h0}^2} + \frac{\log h_{i,t+1}}{w_i} \right) \quad \text{and} \quad \sigma_{ic}^2 = \frac{\sigma_{h0}^2w_i}{\sigma_{h0}^2 + w_i},$$

and the acceptance probability is equal to one since there is no previous draw. For the last element we have

$$\mu_{iT} = \log h_{i,t-1} \quad \text{and} \quad \sigma_{ic}^2 = w_i,$$

where the prior on the distribution of $\log h_0$, providing values for the mean $\mu_{i0}$ and the variance $\sigma_{h0}^2$, is described in Appendix 2.B.

\(^{27}\)In the first iteration, the squared orthogonalized residuals $v_{i,t}^2$ are used to initialize the volatilities, which are calculated by applying the OLS estimates from the initial sample on the actual sample.
(iv) **Drawing hyperparameters.** The hyperparameters of the model are the covariance matrices of the innovations, i.e. $Q$ (coefficient states), $S$ (covariance states), and $W$ (volatility states). Conditional on $y^T$, $\beta^T$, $A^T$, and $H^T$, the state innovations are observable. Since the hyperparameters are assumed to be independent, each covariance matrix can be drawn from its respective distribution. Since we have restricted the hyperparameter matrix $Q$ to be diagonal, its diagonal elements $q_i$ have univariate inverse Gamma distributions with scale parameter $\gamma_{i,1}^q$ and degrees of freedom $\delta_{i,1}^q$:

$$f(q_i|y^T, \beta^T) = IG\left(\frac{\gamma_{i,1}^q}{2}, \frac{\delta_{i,1}^q}{2}\right),$$

where $\delta_{i,1}^q = \delta_{0}^q + T$ and $\gamma_{i,1}^q = \gamma_{i,0}^q + \sum_{t=1}^{T}\epsilon^q_{i,t}$ (see e.g. Kim and Nelson, 1999). Similarly, restricting $S$ to be diagonal, each of its diagonal elements $s_i$ has an inverse Gamma distribution with scale parameter $\gamma_{i,1}^s$ and degrees of freedom $\delta_{i,1}^s$:

$$f(s_i|y^T, A^T) = IG\left(\frac{\gamma_{i,1}^s}{2}, \frac{\delta_{i,1}^s}{2}\right),$$

where $\delta_{i,1}^s = \delta_{0}^s + T$ and $\gamma_{i,1}^s = \gamma_{i,0}^s + \sum_{t=1}^{T}\nu^s_{i,t}$. Finally, the diagonal elements $w_i$ of $W$ have univariate inverse Gamma distributions with scale parameter $\gamma_{i,1}^w$ and degrees of freedom $\delta_{i,1}^w$:

$$f(w_i|y^T, H^T) = IG\left(\frac{\gamma_{i,1}^w}{2}, \frac{\delta_{i,1}^w}{2}\right),$$

where $\delta_{i,1}^w = \delta_{0}^w + T$ and $\gamma_{i,1}^w = \gamma_{i,0}^w + \sum_{t=1}^{T}\omega^w_{i,t}$.

**Summary.** The Gibbs sampling algorithm is summarized as follows:

1. Initialize $R^T$, $Q$, $S$, and $W$.
2. Draw coefficients $\beta^T$ from $f(\beta^T|y^T, R^T, Q)$.
3. Draw covariances $A^T$ from $f(A^T|y^T, H^T, S)$.
4. Draw volatilities $H^T$ from $f(H^T|y^T, \beta^T, A^T, W)$.
5. Draw hyperparameters from $f(q_i|y^T, \beta^T)$, $f(s_i|y^T, A^T)$, and $f(w_i|y^T, H^T)$.
6. Go to step 2.
2.B Calibration of the priors

This appendix discusses the calibration of our priors. We closely follow common choices in the TVP-VAR literature and impose relatively conservative priors, particularly on the hyperparameters (see e.g. Benati and Mumtaz, 2007; Cogley and Sargent, 2002, 2005; Primiceri, 2005).

However, unlike most previous studies those priors are not calibrated based on OLS estimates from an initial training sample which is then discarded. This strategy would force us to sacrifice part of our already relatively short sample. Instead, we calibrate our priors based on OLS estimates from the full sample. This type of strategy is suggested by Canova (2007) and Canova and Ciccarelli (2009) for cases where a training sample is not available. A fixed-coefficient VAR model is thus estimated by OLS (equation by equation) on the full sample from 1980Q1 to 2008Q4.

**VAR coefficients.** Let \( \hat{\beta} \) denote the OLS estimate of the VAR coefficients and \( \hat{\Xi} \) their covariance matrix. We set

\[
\beta_0 \sim N(\hat{\beta}, 4 \times \hat{\Xi}),
\]

where the variance scaling factor increases the uncertainty about the size of the VAR coefficients in the initial sample versus the actual sample.

**Elements of \( H_t \).** Denote the OLS estimate of the VAR covariance matrix as \( \hat{\Sigma} \). We apply a triangular decomposition of this matrix similar to equation (2.4) in the main text, \( \hat{\Sigma} = \hat{\Psi}^{-1}\hat{\Phi}(\hat{\Psi}^{-1})' \), and denote the vector of diagonal elements of \( \hat{\Phi} \) as \( \phi_0 \). Our prior for the diagonal elements of the matrix \( H_t \) is

\[
h_0 \sim N(\phi_0, 10 \times I).
\]

The variance scaling factor 10 is arbitrary but large relative to the mean \( \phi_0 \).
Elements of $A_t$. Denote the vector of non-zero off-diagonal elements of $\hat{\Psi}$ as $\psi_0$, ordered by rows. The prior for the elements of $A_t$ is

$$\alpha_0 \sim N(\psi_0, 10 \times \text{diag}(\psi_0)),$$

where the variance of $\alpha_0$ is scaled up taking into the magnitude of the respective elements of the mean $\psi_0$, as in Benati and Mumtaz (2007).

Hyperparameters. The prior on the diagonal elements of the coefficient state error variance $Q$ is also of the inverse Gamma type:

$$q_i \sim IG\left(\frac{\gamma_{i,0}^q}{2}, \frac{\delta_0^q}{2}\right),$$

where $\gamma_{i,0}^q = k_Q \times \hat{\xi}_i$, where $\hat{\xi}_i$ denotes the $i$-th diagonal element of the OLS covariance matrix $\hat{\Xi}$ and $k_Q = 10^{-4}$. Hence, our prior attributes only 0.01 percent of the uncertainty surrounding the OLS estimates to time variation, following Cogley and Sargent (2002). The degrees of freedom $\delta_0^q$ are set to one, which is the minimum for the prior to be proper. We thus put as little weight on the prior as possible. The prior on the diagonal elements of the hyperparameter matrix $S$ for the covariance states is also of the inverse Gamma type:

$$s_i \sim IG\left(\frac{\gamma_{i,0}^s}{2}, \frac{\delta_0^s}{2}\right),$$

where $\gamma_{i,0}^s = k_S \times \hat{\psi}_i$, where $\hat{\psi}_i$ denotes the $i$-th diagonal element of the OLS covariance matrix $\hat{\Psi}$ and $k_S = 10^{-2}$. Here we follow Primiceri (2005), who makes similar choices for a block diagonal structure of $S$. The degrees of freedom $\delta_0^s$ are again set to the minimum value of one. The prior on the diagonal elements of the variance $W$ for the volatility states is also of the inverse Gamma type:

$$w_i \sim IG\left(\frac{\gamma_{i,0}^w}{2}, \frac{\delta_0^w}{2}\right),$$

where $\gamma_{i,0}^w = k_W$. We set $k_W = 10^{-4}$ and $\delta_0^w = 1$. The parameters of the distribution are the same as in Cogley and Sargent (2005), and Benati and Mumtaz (2007).
2.C Convergence of the Markov chain

This appendix assesses the convergence of the Markov chain produced by the Gibbs sampler. We apply three types of convergence checks to the VAR coefficients, the covariances, and the volatilities. The hyperparameters are omitted in these checks, because they are not the direct objects of interest.

The first convergence check are the diagnostics due to Raftery and Lewis (1992), which are used to assess the total number of iterations required to achieve a certain precision, and the minimum burn-in period and thinning factor. The parameters for the diagnostics are specified as follows: quantile = 0.025; desired accuracy = 0.025; required probability of attaining the required accuracy = 0.95. We generate a Markov chain with 5,000 draws which is then used as the input for the diagnostics as suggested by Raftery and Lewis (1992). Table 2.4 reports the diagnostics. For all three state vectors, the required number of runs is far below the total number of iterations actually applied. The same holds for the number of burn-in replications and the thinning factor. The choices made to generate the Markov chain therefore seem to be validated.

The second convergence check are the inefficiency factors (IFs) for the posterior estimates of the parameters. The IF is the inverse of Geweke’s (1989) relative numerical efficiency measure, i.e. $IF = 1 + 2 \sum_{k=1}^{\infty} \rho_k$, where $\rho_k$ is the $k$-th order autocorrelation of the chain. This diagnostic therefore serves to judge how well the chain mixes. Primiceri (2005) argues that low autocorrelations suggest that the draws are close to independent, which increases the efficiency of the algorithm. We use a 4 percent tapered window for the estimation of the spectral density at frequency zero. Values of the IFs below or around 20 are regarded as satisfactory, according to Primiceri (2005). The left panels of Figure 2.14 report the IFs for the state vectors. The IFs are far below 20 for the coefficients and the covariances, but around 30 to 35 for the volatilities. Compared to the results reported e.g. in Primiceri (2005) and considering the lower number of observations in our sample, however, these results still seem satisfactory.

The final convergence test applied is the convergence diagnostic (CD) due to Geweke (1992). According to Koop (2003), this diagnostic is based on the idea that,

---

28See Koop (2003), Chapter 4, for a review of convergence diagnostics.
Table 2.4: Raftery and Lewis (1992) diagnostics\textsuperscript{a,b}

<table>
<thead>
<tr>
<th>Estim. parameters</th>
<th>Thinning factor</th>
<th>Burn-in replic.</th>
<th>Total runs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficients</td>
<td>4068</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Covariances</td>
<td>452</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Volatilities</td>
<td>678</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The parameters for the Raftery and Lewis (1992) diagnostics are as follows: quantile = 0.025, desired accuracy = 0.025, required probability of attaining the required accuracy = 0.95.

\textsuperscript{b} The results are based on 5,000 iterations of the Gibbs sampler with zero burn-in replications and thinning factor equal to one.

if a sufficiently large number of draws has been taken, the posterior estimates based on the first half of draws should be essentially the same as the estimates based on the second half of draws. If they are very different, either too few draws have been taken and estimates are inaccurate or the effects of the initial values of the chain have not worn off. We therefore divide the 1,000 draws from the posterior distribution into a first set of $N_1 = 100$ draws, a middle set of 500 draws, and a last set of $N_2 = 400$ draws, as suggested by Koop (2003). The middle set of draws is dropped to make it likely that the first set and the last set are independent of each other, which is assessed by the diagnostic. The convergence diagnostic is given by

$$CD = \frac{\hat{\theta}_1 - \hat{\theta}_2}{\hat{\sigma}_1 / \sqrt{N_1} + \hat{\sigma}_2 / \sqrt{N_2}} \rightarrow N(0, 1),$$

by a central limit theorem, where $\hat{\theta}_i$ and $\hat{\sigma}_i / \sqrt{N_i}$ denote the posterior means of the parameters and their numerical standard errors based on the $i$-th set of draws, for $i = 1, 2$ (see Koop, 2003). We plot the $p$-values for the null hypothesis that the two sets of draws are the same in the right panels of Figure 2.14. The $p$-values are mostly larger than conventional significance levels for the VAR coefficients and the covariances, indicating that a sufficiently large number of draws has been taken for these parameters. The null hypothesis is often rejected for the volatilities, but this outcome did not change when a larger number of draws was taken.

To summarize, the coefficients and covariances have in general better convergence properties than the volatilities. Since the focus of our analysis is on impulse responses
Figure 2.14: Convergence diagnostics for state vectors

Notes. Horizontal axes refer to vectors of time-varying parameters with one point representing one parameter at a given time (e.g. volatilities $h_{i,t}$); left panels: inefficiency factors, i.e. inverse of Geweke’s (1992) relative numerical efficiency measure; computed as $IF = 1 + 2 \sum_{k=1}^{\infty} \rho_k$, where $\rho_k$ is the $k$-th order autocorrelation of the Markov chain; right panels: $P$-values of Geweke’s (1992) convergence diagnostic; computed as $CD = (\hat{\theta}_1 - \hat{\theta}_2)/(\hat{\sigma}_1/\sqrt{N_1} + \hat{\sigma}_2/\sqrt{N_2}) \to N(0,1)$, where $N_1 = 100$, $N_2 = 400$, middle 500 draws dropped.

2.D Detailed data description

This appendix provides details on the data definitions used in the main text. Throughout, AWM refers to the Area-Wide Model database (see Fagan et al., 2005), BIS to the Bank of International Settlements macro-economic series, CE to the Consensus Eco-
nomics survey data, EC to the European Commission forecasts and PPP to the dataset provided by Paredes et al. (2009), to which we refer for details on the construction of the fiscal variables. All quarterly series are provided in seasonally adjusted terms from the original sources, except for the HICP of which we take annual differences.

- **Government spending:** Sum of nominal general government final consumption expenditure (variable GCN in PPP) and nominal general government investment (variable GIN in PPP), euro area aggregates, scaled by GDP deflator plus labor force and transformed into natural logarithms.

- **GDP:** Aggregate euro area real gross domestic product, variable YER in the AWM database, where it is calculated as a weighted average of national variables. The original source of GDP and its components in AWM is Eurostat; the variables are then re-scaled to the ECU-euro corrected level of 1995 and backdated with rates of growth of the original AWM series.\(^{29}\)

- **Private consumption:** Aggregate euro area private consumption, constructed by multiplying real private consumption (variable PCR in AWM) with the private consumption deflator (variable PCD in AWM), divided by GDP deflator plus labor force and transformed into natural logarithms.

- **Interest rate:** Weighted euro area short-term nominal interest rate, variable STN in AWM, where it is calculated as a weighted average of national variables taken from the ECB Monthly Bulletin and backdated with the corresponding series contained in the original database (source: Bank of International Settlements and European Commission’s AMECO database).

- **Private investment:** Aggregate euro area total economy gross investment minus general government investment (nominals), scaled by GDP deflator plus labor force and transformed into natural logarithms. Total economy investment corresponds to the variable ITR in AWM, government investment is the variable GIN in PPP.

\(^{29}\)The weights used in AWM are based on constant GDP at market prices for the euro area for 1995.
• **Wage rate**: Nominal hourly wage per head (variable WRN in AWM) divided by GDP deflator. The nominal wage in AWM is calculated as a weighted average of national variables.

• **Net taxes**: Non-interest nominal general government revenue (variable TOR in PPP) minus transfers, which include all expenditure items except government consumption, government investment, and interest payments (variable INP in PPP), scaled by GDP deflator plus labor force and transformed into natural logarithms. The general government primary balance is thus the difference between net taxes and government spending.

• **Inflation rate**: Annual rate of change of the Harmonized Index of Consumer Prices, i.e. variable HICP in AWM, where it is calculated as a weighted average of national variables using 1995 HICP weights.

• **GDP deflator**: Index with base year 1995, variable YED in AWM. Deflators in AWM are taken directly from the corresponding ECB Monthly Bulletin series, which are compiled by ECB staff as a weighted average of the national deflators using purchasing power parity adjusted weights.

• **Labor force**: Total euro area labor force, persons, variable LFN in AWM.\(^{30}\)

• **Debt-to-GDP ratio**: Ratio of the outstanding (end-of-period) aggregate euro area stock of nominal public debt over nominal annual euro area GDP, variable GDN_YEN in AWM.

• **Imports-to-GDP ratio**: Ratio of nominal quarterly aggregate euro area imports (variable MTR times MTD in AWM) over nominal quarterly euro area GDP.

• **Credit to households over GDP**: Outstanding total euro area (end-of-period) stock of bank loans to households, variable BISM.Q.COVA.XM.03 in BIS, divided by the sum of nominal euro area GDP of the last four consecutive quarters.

---

\(^{30}\)The labor force is used as a proxy for total population, since quarterly data on total population is not available from AWM for the entire sample period.
• Government budget deficit-to-GDP ratio, forecast (EC): (Minus) general government balance as percentage of GDP, one-year ahead forecasts by EC published in November of the previous year.Forecasts for the euro area are available from 1999 onwards; for previous years up to 1982, aggregate forecasts are constructed from forecasts for Belgium, France, Germany, Greece, Ireland, Italy, Luxembourg, and the Netherlands by aggregating the individual country series using as weights constant GDP at market prices for 1995.

• GDP growth, forecast (EC): Annual real GDP growth rate, one-year ahead forecasts by EC published in November of the previous year. Forecasts for the euro area are only available from 1999 onwards; for previous years up to 1982, aggregate forecasts are constructed as described above.

• Government budget deficit growth, forecast (CE): Consensus mean forecast of (minus) the general government budget balance, converted into growth rates, minus the Consensus mean forecast of consumer price inflation. Both forecasts are computed as the average of one-year ahead forecasts made in each month of the previous year. Forecasts for the euro area are available from 2003 onwards; for previous years up to 1994, aggregate forecasts are constructed from forecasts for France, Germany, and Italy by aggregating the individual country series using as weights constant GDP at market prices for 1995.

• GDP growth, forecast (CE): Consensus mean forecast of the annual real GDP growth rate, computed as the average of one-year ahead forecasts made in each month of the previous year. Forecasts for the euro area are available from 2003 onwards; for previous years up to 1992, aggregate forecasts are constructed as described above.

• Consumption growth, forecast (CE): Consensus mean forecast of the annual real private consumption growth rate, computed as described above.

• Interest rate, forecast (CE): Consensus mean forecast of the short-term (3-month) nominal interest rate, computed as described above.