ISSN 1745-8587

Birkbeck
UNIVERSITY OF LONDON

School of Economics, Mathematics and Statistics

Birkbeck Working Papers in Economics & Finance

BWPEF 1006

The Effects of Fiscal Shocks in SVAR Models: A Graphical Modelling Approach

Matteo Fragetta
University of Salerno

Giovanni Melina
Birkbeck, University of London

February 2010

▪ Birkbeck, University of London ▪ Malet Street ▪ London ▪ WC1E 7HX ▪

The Effects of Fiscal Shocks in SVAR Models: A Graphical Modelling Approach*

Matteo Fragetta[†] Giovanni Melina[‡]

February 22, 2010

Abstract

We apply graphical modelling theory to identify fiscal policy shocks in SVAR models of the US economy. Unlike other econometric approaches – which achieve identification by relying on potentially contentious *a priori* assumptions – graphical modelling is a data based tool. Our results are in line with Keynesian theoretical models, being also quantitatively similar to those obtained in the recent SVAR literature *à la* Blanchard and Perotti (2002), and contrast with neoclassical real business cycle predictions. Stability checks confirm that our findings are not driven by sample selection.

Keywords: Fiscal policy, SVAR, graphical modelling
JEL classification: E62, C32

*We are grateful to Yunus Aksoy, Sergio Destefanis, John Driffill, Marco Reale, Lucio Sarno, Ron Smith, Granville Wilson and seminar participants at Birkbeck College and Salerno University for helpful comments. Moreover, we thank Marco Reale and Granville Wilson for kindly sending some of their Matlab codes, and the developers of JMulTi for providing their software as an open source. The usual disclaimer applies.

[†]Department of Economics, University of Salerno, Via Ponte don Melillo, 84084 Fisciano (SA), Italy; Tel: +39-089-405232 ; fax: +39-089-405232; *e-mail*: mfragetta@unisa.it

[‡]Department of Economics, Birkbeck College, University of London, Malet Street, London WC1E 7HX, United Kingdom; Tel: +44-20-7631-6426; fax: +44-20-7631-6416; *e-mail*: g.melina@ems.bbk.ac.uk

1 Introduction

In macroeconomics there is still no widespread consensus on the impact and transmission channels of fiscal policy on many variables. Both theoretical and empirical studies generally find a positive response of output and hours worked to a positive shock to government purchases and the disagreement is usually about the magnitude and timing of the response. On the contrary, the sign of the responses of variables such as consumption, wages and investment is still a matter of debate.

In the theoretical literature, on one hand neoclassical Real Business Cycle (RBC) theory claims that a positive government spending shock triggers a negative wealth effect that dampens consumption, fosters labour supply and curbs real wages (e.g. Baxter and King (1993)). On the other hand, Keynesian theories and recent dynamic stochastic general equilibrium (DSGE) models such as Linnemann (2006), Ravn et al. (2006) and Gali et al. (2007) among others, assert that an expansionary fiscal policy boosts consumption, hours worked and real wages. In addition, while RBC theories predict that real output should rise less than proportionally to the increase in government spending, due to the crowding-out effect on consumption, Keynesian theories foresee that the increase in consumption should amplify the expansionary effect on output.

The empirical literature of the late 1990s, such as Ramey and Shapiro (1998) and Edelberg et al. (1999), mostly relying on vector-autoregressions (VAR) employing a narrative approach to identify discretionary fiscal policy shocks, supports RBC predictions. More recent empirical studies, starting from Blanchard and Perotti (2002), adopt structural VARs (SVAR) for the purpose of identification and obtain results more in line with Keynesian claims.

Indeed, VAR analysis is a standard tool to understand what happens in actual economies and to evaluate competing theoretical economic models. SVARs, however, generally require the imposition of a number of restrictions, which is often a complex and contentious task, as they may be based on possibly arguable assumptions.

In this paper we conduct a SVAR analysis for the US economy that combines Graphical Modelling (GM) theory. Such an approach allows us to obtain identifying restrictions from statistical properties of the data. The starting point is the computation of partial correlations among the variables in the model and the subsequent construction of a Conditional Independence Graph (CIG), a graphical representation of all statistically significant interconnections among all variables. From the CIG, based on well defined statistical rules, we derive Directed Acyclic Graphs (DAGs), graphical representations of the many possible structural VARs, which are later evaluated by means of statistical information criteria.

Our results are generally in line with the recent SVAR literature (Blanchard and Perotti (2002), Perotti (2005), Caldara and Kamps (2008) among others) and hence give credit to Keynesian claims. In response to a positive government spending shock, we detect a partially deficit-financed fiscal policy and obtain a fiscal multiplier of output greater than one. Adding more recent data increases the magnitude of the fiscal multiplier compared to earlier studies such as Blanchard and Perotti (2002). Private consumption shows a positive and persistent response to a spending shock. While non-residential investment is significantly crowded out by the fiscal expansion, residential investment rises, comoving with output. However, the crowding-out effect on non-residential investment is not stable over the sample considered. Lastly, a positive response of real wages coexists with an increase in hours worked. As far as the effects of a positive tax shock are concerned, we find that peak responses of consumption, non-residential investment and the initial response of hours worked show signs consistent with a negative wealth effect. Subsample checks confirm that results are stable over the sample.

The remainder of the paper is structured as follows. Section 2 provides an overview of recent theoretical and empirical contributions in the field of the macroeconomic implications of fiscal policy. Section 3 illustrates the principles of graphical modelling and how it can be used to identify SVARs. Section 4 presents the data. Section 5 identifies a number of SVARs to evaluate the effects of fiscal policy shocks on the US economy. Section 6 describes the econometric results and conduct some stability checks. Finally, Section 7 concludes.

2 Literature review

Following a positive shock to government spending, textbook IS-LM theory predicts that consumption should rise and thus amplify the expansionary effects of spending on output. In constrast, as shown by Baxter and King (1993), neoclassical Real Business Cycle (RBC) theory generally predicts a positive response of investment and a negative response of consumption and wages. As Galí et al. (2007) point out, this substantial difference across the two classes of models lays on the more or less implicit assumption made on the behaviour of consumers: in the IS-LM model, consumption only depends on current disposable income, hence consumers are all non-Ricardian; in the RBC model consumption depends on life-time wealth, hence consumers are all optimising Ricardian agents. In the RBC model, an increase in government purchases, through an increase in current and/or future taxes, triggers a negative wealth effect that decreases consumption, increases labour supply and decreases wages. The increase in the marginal product of capital, allowed by the increase in labour, causes also a positive reaction of investment.

The early empirical literature, mainly using vector-autoregressions (VAR), supports the RBC claims. In general, empirical studies aiming at studying the effects of fiscal policy shocks confront great difficulties in identifying such shocks, as they have to disentangle the role of automatic stabilisers responding to business cycles from the effects of discretionary fiscal policies. Ramey and Shapiro (1998) introduce the dummy-variable or narrative approach, due to Hamilton (1985), in the context of fiscal policy, though in a univariate setting. The methodology consists in constructing a dummy variable that takes value one at quarters when large military build-ups took place in the US, in order to identify episodes of discretionary fiscal policy. Edelberg et al. (1999) extend this methodology to a multivariate context, and Burnside et al. (2000),[1] as well as Eichenbaum and Fisher (2005) make some modifications. Despite slight methodological differences, all these studies generally reach the same conclusions, at least from a qualitative point of view: in response to a discretionary substantial positive government spending shock, output increases, consumption and wages decline, non-residential investment rises, while residential investment falls. Therefore, these findings support the neoclassical business cycle literature.

More recent empirical studies aiming at detecting the effects of government spending shocks make use of structural vector-autoregressions in order to trace the impulse responses of the macroeconomic variables of interest. Fatás and Mihov (2001) find that there is a strong, positive and persistent impact of fiscal expansions on economic activity. Blanchard and Perotti (2002) achieve identification by relying on institutional information about the tax collection, constructing the automatic response of fiscal variables to the business cycle and, by implication, identifying discretionary fiscal policy shocks. The Blanchard-Perotti approach yields a positive effect of a government spending shock on output and consumption and a negative effect on investment. While these findings are perfectly reasonable in a Keynesian world, they are difficult to reconcile with the RBC literature. Perotti (2005) extends the structural VAR methodology to other countries and reaches similar conclusions. Moreover, Perotti (2007) proposes a variant of the narrative approach that allows the responses to each Ramey-Shapiro episode to have both a different intensity and a different shape. In addition, the author introduces a different method to build the dummy variable, which allows to isolate the abnormal fiscal events and to estimate the normal dynamic response of the non-fiscal variables to these events. Using this methodology the response of consumption is positive, in line with the structural VAR approach. Mountford and Uhlig (2008) extends Uhlig (2005)'s sign-restriction approach to fiscal policy and find a negative response of investment to a fiscal

[1] Within the field of RBC models, Edelberg et al. (1999), in addition to their econometric analysis, also build a variant of the neoclassical growth model distinguishing between residential and non-residential investment to match their empirical findings. Instead, Burnside et al. (2000) introduce habit formation in consumption and adjustment costs in investment to better mimic the timing and quantitative responses of hours worked, investment and consumption.

expansion. However, they find a small response of consumption, significant only on impact.

In order to match the most recent empirical findings, the theoretical literature has recently worked on models able to explain the positive response of consumption to a fiscal expansion. Linneman (2006) uses a non-additively separable utility function within a neoclassical growth model that is able to mimic a pattern for consumption similar to the one found by the structural VAR literature. Ravn et al. (2006) assume habit persistence on the consumption of individual differentiated goods, which implies a countercyclical mark-up of price over the marginal cost. A government spending shock has a negative wealth effect, yet it also boosts aggregate demand, firms reduce their mark-up, labour demand increases and offsets the negative income effect affecting labour supply. As a result, wages and consumption rise. Galí et al. (2007) cast fiscal policy into a new-Keynesian sticky-price model modified to allow for the presence of rule-of-thumb behaviour. Non-Ricardian households are key for the purposes of the model, as they partly insulate aggregate demand from the negative wealth effects generated by the higher levels of current and future taxes needed to finance the fiscal expansion. In this model, the magnitudes of the responses of output and consumption are systematically greater than those generated by the neoclassical model. In addition, the increase in aggregate hours coexists with an increase in real wages. This response of wages is made possible by sticky prices. In fact, even in the face of a drop in the marginal product of labour, real wages can increase as the price mark-up may adjust sufficiently downward to absorb the resulting gap. The combined effect of higher real wages and higher employment raises labour income and stimulates consumption of rule-of-thumb households, and the overall effect on consumption is positive.

Table 1 summarises the results of the theoretical and empirical literature surveyed.

3 Econometric methodology

This section outlines the econometric methodology that we employ to analyze the effects of fiscal policy shocks. Subsection 3.1 explains the basic principles of graphical modelling. Subsection 3.2 illustrates how graphical modelling can be used to identify structural shocks in a VAR framework.

3.1 Graphical modelling

Graphical modelling is a statistical approach aiming at uncovering statistical causality from partial correlations observed in the data, which can be interpreted as linear predictability in the case of a linear regression model.[2]

[2] In this context maximum likelihood estimation would be equivalent.

Primal contributions to the methodology are due to Dempster (1972) and Darroch et al. (1980).

The initial step of the procedure is to compute partial correlations between variables, the significance of which can be tested by using appropriate statistics. Statistically significant partial correlations can be then represented by an undirected graph called *Conditional Independence Graph* (CIG), where random variables are represented by nodes and a significant partial correlation between any two random variables – conditioned on all the remaining variables of the model – represented by a line known as *undirected edge*. In Figure 1.A, we show an example of a CIG. For instance, the edge connecting nodes A and B represents a significant partial correlation between A and B conditioned on C. A significant partial correlation implies conditional dependence if the variables are jointly distributed as a multivariate Gaussian distribution, hence the name Conditional Independence Graph.

When an arrow links the nodes of a CIG, we obtain what is called *Directed Acyclic Graph* (DAG). DAGs and CIGs imply a different definition of joint probability. For example, if we consider Figure 1.A, we can assert that A and C are independent, conditional on C. Therefore, the joint distribution implied by the CIG is the following:

$$f_{A,C|B} = f_{A|B} f_{C|B}$$

while a corresponding DAG such as the one in Figure 1.C2 has a joint distribution equal to:

$$f_{A,B,C} = f_{C|B} f_{B|A} f_A.$$

Nevertheless, there is a correspondence between the two, represented by the so-called *moralization rule*, as firstly shown by Lauritzen and Spiegelhalter (1988). In fact, there is always a unique CIG deriving from a given DAG, obtained by transforming arrows into undirected edges and linking unlinked *parents* of a *common child* with a *moral edge*. In the DAG shown in Figure 1.B1, A and C are parents of B. In order to obtain the corresponding unique CIG we must transform arrows into edges and add a moral edge between parents A and C as in Figure 1.B2. Statistically, when both A and C determine B, a significant partial correlation due to moralization should be observed between A and C.[3]

While there is a unique CIG deriving from a given DAG, the reverse is not true. What we can observe in the data is a CIG, where every edge can

[3]An example should provide a more intuitive insight into the moralization rule: if one wants to become a famous football player (P), he/she must have good skills (S) and/or must work hard (W). Therefore S and W are determinants of P. Conditional on P, there may be cases where S is high and W is low; cases where W is high and S is low; and cases where both S and W are high. There cannot be cases where S and W are both low, otherwise we would not observe P. This example shows that S and W are (negatively) correlated given P.

assume two possible directions. Therefore, for any given CIG, there are 2^n hypothetical DAGs, where n is the number of edges. Figure 1.C shows all the hypothetical DAGs corresponding to the CIG in Figure 1.A. According to what we have said above, the DAG in Figure 1.C1 is not compatible with the CIG, because the moralization rule requires a moral edge between A and C, which is not captured by the CIG.

In the process of obtaining plausible DAGs from an observed CIG, it might also be possible that some of the links captured by the CIG are due to moralization and hence must be eliminated in a corresponding DAG. Such *demoralization* process, in most cases, can be assessed by considering some quantitative rules. Let us suppose we observe a CIG such as the one in Figure 1.B2. If the true corresponding DAG were the one in Figure 1.B1, then the partial correlation between A and C, $\rho_{A,C|B}$, should be equal to $-\rho_{A,B|C} \times \rho_{B,C|A}$. In such a case, when tracing DAG 1.B1, the edge between A and C must be removed.

Any DAG, by definition, has to satisfy the principle of acyclicality. Therefore, the graph depicted in Figure 2 cannot be a DAG as it is clearly cyclic. The acyclicality in a DAG allows to completely determine the distribution of a set of variables and implies a recursive ordering of the variables, where each element in turn depends on none, one or more elements. For example, in the DAG in Figure 1.C2, A depends on no other variables, B depends on A and C on B.

3.2 Graphical modelling in the identification of a SVAR

Graphical Modelling (GM) theory can be applied to obtain identification of structural VARs (SVAR), as shown by Reale and Wilson (2001) and Oxley et al. (2009) among others. This literature considers GM as a data-driven approach that represents a possible solution to the problem of imposing restrictions to identify a SVAR.

Any SVAR may be turned into a DAG where current and lagged variables are represented by nodes and causal dependence by arrows. To do so, we need to establish pairwise relationships among contemporaneous variables in terms of partial correlations conditioned on all the remaining contemporaneous and lagged values. In many cases, it is possible to obtain more parsimonious models since some lagged variables do not play any significant role in explaining contemporaneous variables and the corresponding coefficient vectors present some zeros.[4] In this paper, however, we will consider SVARs where the data generating process presents all the lagged values, as it is standard practice in the applied econometric literature aiming at analyzing the impulse responses of a set of macroeconomic variables. The first step in constructing a DAG representation of a SVAR is the determination

[4]Reale and Tunnicliffe (2001) and Oxley et al. (2009) argue that, in some cases, a sparse lag structure may yield models with better statistical properties.

of the lag order through the minimization of an order selection criterion such as the Akaike Information Criterion (AIC), the Hannan and Quinn Information Criterion (HIC) or the Schwarz Information Criterion (SIC). We can then derive a pth-order vector autoregressive model, m-dimensional time series $X_t = (x_{t,1}, x_{t,2}, ..., x_{t,m})$ in canonical (or reduced) form, which can be expressed as:

$$X_t = c + A_1 X_{t-1} + A_2 X_{t-2} + ... + A_p X_{t-p} + u_t$$

where c allows for a non-zero mean of X_t, each variable is expressed as a linear function of its own past values, the past values of all other variables being considered and a serially uncorrelated innovation u_t, whose covariance matrix V is generally not diagonal. The correlation between two errors represents the partial correlation of two contemporaneous variables conditioned on all the lagged values. Hence, in order to construct the CIG among contemporaneous variables conditioned on all the remaining contemporaneous and lagged variables, we can derive the sample partial correlation between the innovations, conditioned on the remaining innovations of the canonical VAR,[5] calculated from the inverse $\hat{W}$ of the sample covariance matrix $\hat{V}$ of the whole set of innovations as:

$$\hat{\rho}(u_{i,t}, u_{j,t} | \{u_{k,t}\}) = -\hat{W}_{ij} / \sqrt{(\hat{W}_{ii} \hat{W}_{jj})}$$

where $\{u_{k,t}\}$ is the whole set of innovations excluding the two considered.

The critical value utilised to test for the significance of the sample partial correlations can be calculated by using the relationship between a regression t-value and the sample partial correlation, as shown by Greene (2003), and considering the asymptotic normal distribution of the t-value for time series regression coefficients. This is given by:

$$\frac{z}{\sqrt{(z^2 + \nu)}} \approx \frac{z}{\sqrt{n - p}}$$

where n is the sample size, $\nu = n - k - 1$ are the residual degrees of freedom obtained as a regression of one variable on all the remaining variables and z represents a critical value at a chosen significance level of the standard normal distribution. Whenever a sample partial correlation is greater than the calculated critical value, a link is retained.

All arrows end in nodes representing contemporaneous variables. At this point, the only causality we can assume is the relationship between lagged and contemporaneous variables determined by the flow of time. Next, we need to consider all the possible DAGs representing alternative competitive

[5] Granger and Swanson (1997) have applied a similar strategy to sort out causal flows among contemporaneous variables, i.e. applying a residual orthogonalization of the innovations from a canonical VAR.

models of the relationships among contemporaneous variables. Finally, we compare the DAGs compatible with the estimated CIG by using likelihood based methods,[6] such as AIC, HIC and SIC, and choose the best-performing DAG.[7]

In order to construct an empirically well-founded SVAR, we have to assure that the covariance matrix of the resulting residuals is diagonal. A first diagnostic check is thus inspecting the significance of such correlations. Further diagnostic checks are possible. For instance, as this procedure typically imposes over-identifying restrictions, a χ^2 likelihood-ratio test can be conducted.

4 Data

In order to make our results comparable with the previous literature, we use the same sample period and data sources as Caldara and Kamps (2008). Therefore, we use quarterly US data over the period 1955:1-2006:4. All series are seasonally adjusted by the source.

Our baseline model is a three-variable VAR that includes the log of real per capita government spending (g_t), the log of real per capita net taxes (t_t) and the log of real per capita GDP (y_t). Government spending and taxes are net of social transfers. Government spending is the sum of government consumption and investment, while net taxes are obtained as government current receipts less current transfers and interest payments.[8] To assess the effects of fiscal policy shocks on a set of key macroeconomic variables, we follow and Blanchard and Perotti (2002) specify four-variable VAR models by adding one variable at a time to the baseline model. The other variables are the log of real per capita private consumption (c_t), the log of per capita hours worked (h_t), the log of the real wage (w_t), the log of real per capita private residential investment (R), and the log of real per capita private non-residential investment (NR).

We extracted the components of GDP, government receipts, and the GDP deflator from the NIPA tables of the Bureau of Economic Analysis. We obtained real hourly compensation from the Bureau of Labor Statistics and

[6] In some cases, the distributional properties of the variables for different DAGs are likelihood equivalent, although the residual series are different. In such cases, it is possible to construct DAG models by considering only the lagged variables that play a significant role in explaining contemporaneous variables determined by the significant partial correlation. This can help, via comparison of information criteria, determine the best DAG for contemporaneous variables.

[7] Even in the presence of non-stationary variables, the sampling properties of GM and the outlined procedure are still valid, as shown by Wilson and Reale (2008).

[8] We converted the components of national income and net taxes into real per-capita terms by dividing their nominal values by the GDP deflator and the civilian population. The latter is available in the ALFRED database of the Federal Reserve Bank of Saint Louis.

the measure of per capita hours worked used in Francis and Ramey (2005) from Ramey's webpage.

5 Identification of structural vector-autoregressions

We study the effects of fiscal policy shocks from a macroeconomic perspective by means of structural VAR models identified through DAGs.

After collecting the endogenous variables of interest in the k-dimensional vector X_t, the reduced-form VAR model associated to it can be written as:

$$X_t = A(L)X_{t-1} + u_t \tag{1}$$

where $A(L)$ is a polynomial in the lag operator L and u_t is a k-dimensional vector of reduced-form disturbances with $E[u_t] = 0$ and $E[u_t u_t'] = \Sigma_u$.[9]

As the reduced-form disturbances are correlated, in order to identify fiscal policy shocks, we need to transform the reduced-form models into structural models. Pre-multiplying both sides of equation (1) by the $(k \times k)$ matrix A_0, yields the structural form:

$$A_0 X_t = A_0 A(L) X_{t-1} + B e_t \tag{2}$$

In our benchmark case we also include a constant and a linear trend among regressors. The relationship between the structural disturbances e_t and the reduced-form disturbances u_t is described by the following:

$$A_0 u_t = B e_t \tag{3}$$

where A_0 also describes the contemporaneous relation among the endogenous variables and B is a $(k \times k)$ matrix. In the structural model, disturbances are assumed to be uncorrelated with each other. In other words, the covariance matrix of the structural disturbances Σ_e is diagonal.

As it is, the model described by equation 2 is not identified. Therefore, first we restrict matrix B to be a $(k \times k)$ diagonal matrix. As a result, the diagonal elements of B will represent estimated standard deviations of the structural shocks. In order to impose identifying restrictions on matrix A_0, we apply *graphical modeling* theory and trace DAGs of the reduced-form residuals.

A feature of DAGs is acyclicality, which implies a recursive ordering of the variables that makes A_0 a lower-triangular matrix. A_0 has generally zero elements also in its lower triangular part, hence, in general, the model is over-identified. The GM methodology has the distinctive feature that the

[9] We report results obtained by using a 4-th order lag polynomial for all models, as it is the usual choice with quartely data and is in line with the related literature. However, using the number of lags suggested by information criteria yields no differences, as we obtain CIGs with the same edges.

variable ordering and any further restrictions come from statistical properties of the data.

Consistently with the methodology described in Section 3, we build DAGs of the residuals obtained by fitting the various specifications to equation (1). In table 2 we report the estimated partial correlation matrices of the series innovations and their significance at 0.10, 0.05 and 0.01 levels. These allows us to draw the CIGs reported in the left column of Figure 3. The statistical strength of the links is represented by dashed, thin or thick lines, which reflect significance at the 0.10, 0.05 and 0.01 levels, respectively.

Applying the GM procedure allows us to define the DAGs reported in the right column of Figure 3.

DAG a2- baseline. The two edges in CIG a1 cannot be moral, as moral edges link parents of a common child. The four possible DAGs implied by CIG *a1* are reported in Figure 4. DAG (B) can be discarded because a moral edge between u_t^g and u_t^t is not captured in the CIG. Hence, we need to compare the three remaining models. Table 3 shows that the three information criteria reported are minimised by the model implied by DAG (A). The best performing DAG implies that government spending is not affected contemporaneously by shocks originating in the private sector. As we employ quarterly data and definitions of fiscal variables that exclude most of the automatic stabilizers, this finding makes economic sense in the light of the typical decision and implementation lags present in the budgeting process. Such an argument is shared by virtually all other related empirical studies. However, while the related literature uses this argument as an *a-priori* identifying assumption,[10] in this paper we obtain it as a result. If we fit DAG (A) to the estimated residuals, we get significant coefficients (t-statistics are reported adjacent to directed edges) and signs compatible with economic arguments. An increase in government spending has a contemporaneous (within a quarter) effect on real output, the tax base increases and, thus, tax receipts contemporaneously rise. As a diagnostic check, we inspect the cross-correlations matrix of the resulting residuals in Table 4 and find that all cross-correlations lie within two standard errors from zero. We use the directions obtained for the baseline variables also in the DAGs that follow.

DAG b2- consumption. Same arguments apply to baseline variables. In addition, $u_t^y \rightarrow u_t^c$, as the opposite would imply a moral link between u_t^g and u_c^t, which does not appear in CIG b1. Fitting this DAG yields significant co-

[10] First, Blanchard and Perotti (2002) argue that in contrast to monetary policy, decision and implementation lags in fiscal policy imply that, within a quarter, there is little discretionary response of fiscal policy to unexpected contemporaneous movements in activity. Next, Caldara and Kamps (2008), when they apply a recursive approach *à la* Choleski, order government spending first. Last, Mountford and Uhlig (2008), when using the sign-restriction approach, define a business cycle shock as a shock which jointly moves output, consumption, non-residential investment and government revenue, but not government spending.

efficients and signs are compatible with economic arguments. In particular, a positive shock to income has a contemporaneous positive effect on consumption. All cross-correlations of the resulting residuals are insignificant.

DAG c2- hours worked. Our best DAG selected on the basis of the information criteria (not reported) indicates a strongly significant coefficient for the contemporaneous output in the hours equation. Moreover an increase in contemporaneous hours worked has a contemporaneous positive effect on tax receipts. The resulting cross correlation between the residual series are all not statistically different from zero.

DAG d2- real wage. The link between u_t^g and u_t^w, significant at a 0.10 level, may be a moral link in the case in which $u_t^w \rightarrow u_t^y$. However, information criteria suggest that $u_t^y \rightarrow u_t^w$. The positive coefficient of u_t^y in the regression of u_t^w captures a positive contemporaneous effect on real wages of a shock to economic activity.

DAG e2- residential investment. We apply analogous arguments to those applied to DAG b2. Here $u_t^y \rightarrow u_t^R$, as the opposite would imply a moral link between u_t^g and u_c^R, which does not appear in CIG e1. Fitting this DAG yields significant coefficients and signs are compatible with economic arguments. In particular, a boom in economic activity has a contemporaneous positive effect on residential investment. All cross-correlations of the resulting residuals are insignificant.

DAG f2- non-residential investment. The CIG for non residential investment is the one with the richest set of contemporaneous relationships among variables. In addition to the relationships already established for the baseline variables, government spending has a contemporaneous (negative) effect on non-residential investment. Therefore we need to establish the relationship between taxes and non residential investment and between output and non residential investment. This makes $2^2 = 4$ potential DAGs. The information criteria select the model in which output has a contemporaneous positive effect on non-residential investment and the latter has a contemporaneous positive effect on tax receipts. The resulting cross correlations between the residuals do not differ statistically from zero.

Now, we can use the DAGs depicted in Figure 3 to impose restrictions on matrix A_0. This allows us to identify our structural VAR models. For illustrative purposes we show what the relationship between the structural disturbances e_t and the reduced-form disturbances u_t looks like in the model for private consumption:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ 0 & -a_{42} & 0 & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^t \\ u_t^c \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^t \\ e_t^c \end{bmatrix}}_{e_t} \tag{4}$$

In the appendix, we report matrices for all models.

As anticipated above, matrix A_0 is over-identified, as the assumption of orthonormal structural innovations imposes $k(k+1)/2$ restrictions on the k^2 unknown elements of A_0, where k is the number of endogenous variables. In the case of four variables, this makes six restrictions. It follows that we have three over-identifying restrictions.

In Table 5, for all models we report the variable ordering in vector X_t, the maximum-likelihood estimates of matrices A_0 and B, and the likelihood-ratio (LR) test for over-identification. All estimated coefficients have the right signs and are significant at least at a 0.05 level. Moreover, we fail to reject LR-tests for over-identification of all models at any reasonable level of significance.

All the estimated SVAR models identify identical structural fiscal policy shocks. This is clearly depicted in Figure 5, where the identified spending and tax shock deriving from the six models above are coincident series.

6 Results

In this section we present empirical results for government spending and tax shocks by analyzing the impulse responses obtained from the SVARs identified above.

Following the procedure by Blanchard and Perotti (2002), we transform the impulse responses of output and its components in such a way that they can be interpreted as multipliers. In other words, they represent the dollar response of the respective variable when the economy is hit by a fiscal shock of size one dollar. To achieve this, first, we divide the original impulse responses by the standard deviation of the respective fiscal shock. This allows us to deal with shocks of size one percent. Second, we multiply the resulting responses by the ratio of the responding variable to the shocked fiscal variable, evaluated at the sample mean.

As far as the impulse responses of hours worked and wages are concerned, we simply express them as percentage-point changes subsequent to a fiscal shock of size one percent. For each variable we report responses for a 40-quarter horizon and 90 percent confidence intervals obtained by applying the procedure due to Hall (1992) with 2000 boostrap replications.

6.1 Government spending shock

In Figure 6 we report results for a government spending shock of one dollar.

Government spending reacts strongly and persistently to its own shock. It reaches its peak of 1.15 dollars after one year and persistently stays above baseline (more 95 percent of the shock is still present after two years).

Real output increases on impact by almost 1.10 dollars, slightly decreases after two quarters, and then rises again up to a peak of 1.75 dollars two years after the shock. Then, it persistently and significantly stays above baseline.

The spending multiplier is greater than one both on impact and at a longer time horizon.

Taxes partly offset the one-dollar increase in government spending, since they rise up to 30 cents, probably as an automatic response to the increase in output (note that the shape of the tax response mimic that of output). Taxes reach their peak of more than 50 cents four quarters after the shock. This suggests that the fiscal expansion is at least partially deficit-financed. The response of taxes is also long lasting and statistically significant on impact, after a year, and at longer horizons.

Private consumption shows a positive, smooth and hump-shaped response to the government spending shock. It increases on impact by 35 cents to reach a peak of almost 90 cents ten quarters after the shock has occurred. Then, it persistently stays above baseline.

Private non-residential investment does not move on impact, but declines afterwards showing a peak crowding-out effect of almost -75 cents one year after the spending shock.

Residential investment reacts positively on impact to the fiscal expansion, probably following the increase in output, rising by almost 10 cents and reaching a peak of 20 cents after one year.

Hours worked react positively to a fiscal expansion of one percent from baseline, rising by 0.10 percent on impact and by 0.14 percent after three quarters. Even if the response of hours is not statistically significant at all quarters, significance is achieved at the mentioned quarters and at a longer horizon.

The response of real wages is slightly negative on impact, as they fall by 0.06 percent. They turn positive after a quarter reaching a peak of 0.20 percent after two years and persistently stay above baseline.

6.2 Tax shock

In Figure 7 we report results for a tax shock of size one dollar.

The tax response reaches its peak on impact and then declines quite smoothly till dying out. The policy experiment shows that a tax increase does not yield any statistically significant effect on output, residential investment, and real wages. Instead, the peak responses of consumption, non-residential investment and hours worked are statistically significant. While consumption decreases by 10 cents two quarters after the shock, non-residential investment rises up to 7 cents three quarters ahead. Last, hours worked reach their peak at 3 percent, three quarters after the tax shock.

As far as the tax policy shock is concerned, from a statistical point of view, we are able to comment only on the peak responses of the mentioned variables. Nevertheless, from an economic point of view, these results are sufficient to detect that a negative wealth effect affects the US economy when the latter is hit by a positive tax shock. In fact, as a consequence of a negative wealth effect we would expect consumption and leisure to decline, i.e. hours worked to rise and private non-residential investment to increase, given the increase in the marginal product of capital determined by the increase in employment. The signs of the peak responses of consumption, non-residential investment and hours worked are consistent with this transmission mechanism.

6.3 Subsample stability

We first employ forecast Chow tests to check the overall stability of the parameters of the estimated models. Once the sample has been split into two parts, this test allows us to detect whether a structural change has occurred, by comparing the full sample residual variance with the residual variance of the whole sample. In other words, the test checks whether forecasts made exploiting the first subsample are compatible with the observations contained in the second subsample. We start from 1961:3 and recursively repeat the test at each subsequent data point. Given the tendency of the test to over-reject the null, i.e. to yield a high type I error (Lütkephol, 2005), we recover p-values with a procedure based on 2000 bootstrap replications. As we observe in Figure 8, the test fails to reject the null of parameter constancy on every occasion at any reasonable significance level.

Then we replicate SVAR estimation and impulse response analysis by removing ten years of observations at a time.

In Figure 9, we report responses to a government spending shock. For all responding variables but non-residential investment, subsample variability does not produce changes in the impulse responses able to controvert the main findings outlined above.

Figure 10 depicts responses to a tax shock. Except for the responses of government spending and real wages, also in this case, removing a decade of observations at a time does not yield very different responses compared to the ones obtained by exploiting the full sample.

In table 6, we report peak responses and their significance. As far as the government spending shock is concerned, the peak responses of all variables except non-residential investment, systematically show the same sign across subsample even if their magnitude varies over time.[11] In particular, removing the last ten years in the sample decreases the fiscal multiplier.

Signs of the peak responses obtained by exerting a positive tax shock are stable for tax revenue itself, consumption, residential and non-residential investment. Hours worked show negative, though insignificant, peak responses. However, for at least two quarters after the shock, responses are positive and statistically significant.

[11] Unlike in the full sample, taxes and hours worked show negative peak responses when the decade 1965:1-1974:4 is removed but these are not statistically significant.

6.4 Relation with other studies

As discussed in Section 2, the recent DSGE literature regards as a stylized fact that private consumption increases when the economy is hit by an expansionary fiscal spending shock. Our empirical results for the US economy, relying on an alternative identification approach to those commonly employed in the related literature, support this claim.

With respect to Blanchard and Perotti (2002), the fiscal multiplier on output is greater: 1.75 against 1.29. This difference depends on the inclusion of more recent data. In fact, when we remove data from 1995:1 to 2006:4, the peak multiplier declines to 1.31.

Consistently with new-generation DSGE models, we also find that while non-residential investment falls, hours worked and real wages rise as a consequence of a positive government spending shock. As in Caldara and Kamps (2008), our results also show quite persistent impulse responses for consumption and real wages. This is not the case in theoretical models such as Gali et al. (2007) where the responses of consumption and wages are initially positive but they turn negative after one year.

As far as the tax shock is concerned, in principle one may argue that real output does not respond on impact to tax shocks because the graphical modelling approach imposes a unique contemporaneous relation from output to taxes, when the contemporaneous effect of taxes on output may be conceptually important. Caldara and Kamps (2008) in applying the Blanchard-Perotti methodology to their data find that the impact response of output to taxes does not significantly differ from zero, which is also captured by our DAGs. Moreover this result is in line with the results reported by Perotti (2005).

7 Conclusion

We have applied graphical modelling theory to identify fiscal policy shocks in SVAR models of the US economy. This approach has allowed us to rely on statistical properties of the data for the purpose of identification.

In response to a positive government spending shock we obtain results in line with Keynesian views. First, real output responds positively and more than proportionally. Next, private consumption shows a positive and persistent response to a spending shock. While non-residential investment is significantly crowded out by the fiscal expansion, residential investment rises, comoving with output. Last, a positive response of real wages coexists with an increase in hours worked.

When we analyse the effects of a positive tax shock, in general, we do not obtain statistically significant impulse responses. However, peak responses of consumption and non-residential investment, as well as the initial response of hours worked are statistically significant and their signs are consistent with a negative wealth effect incepted by the increase in taxation.

The outlined results are stable over the sample period. In general, adding more recent data increases the magnitude of the fiscal multiplier compared to earlier studies. The crowding-out effect on non-residential investment is not systematically captured in all subsamples.

References

Baxter, M. and King, R. (1993). Fiscal policy in general equilibrium. *The American Economic Review*, pages 315–334.

Blanchard, O. and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics*, 117(4):1329–1368.

Burnside, C., Eichenbaum, M., and Fisher, J. (2000). Assessing the effects of fiscal shocks. *NBER Working Papers*, 7459.

Caldara, D. and Kamps, C. (2008). What are the effects of fiscal shocks? A var-based comparative analysis. *ECB Working Paper Series*, 877.

Darroch, J., Lauritzen, S., and Speed, T. (1980). Markov fields and log-linear interaction models for contingency tables. *The Annals of Statistics*, pages 522–539.

Dempster, A. (1972). Covariance selection. *Biometrics*, 28(1):157–175.

Edelberg, W., Eichenbaum, M., and Fisher, J. (1999). Understanding the Effects of a Shock to Government Purchases. *Review of Economics Dynamics*, pages 166–206.

Eichenbaum, M. and Fisher, J. D. M. (2005). Fiscal policy in the aftermath of 9/11. *Journal of Money, Credit and Banking*, 37(1):1–22.

Fatás, A. and Mihov, I. (2001). The effects of fiscal policy on consumption and employment: theory and evidence. *CEPR Discussion Papers*, 2760.

Francis, N. and Ramey, V. (2005). Measures of per capita hours and their implications for the technology-hours debate. *NBER working paper*.

Galí, J., López-Salido, J., and Valles, J. (2007). Understanding the effects of government spending on consumption. *Journal of the European Economic Association*, 5(1):227–270.

Granger, C. and Swanson, N. (1997). Impulse response functions based on a casual approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92(1):357–367.

Greene, W. (2003). *Econometric analysis.* Prentice Hall Upper Saddle River, NJ.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag: New York.

Hamilton, J. (1985). Historical causes of postwar oil shocks and recessions. *The Energy Journal*, 97–116.

Lauritzen, S. and Spiegelhalter, D. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 157–224.

Linneman, L. (2006). The effect of government spending on private consumption: A puzzle? *Journal of Money Credit and Banking*, 38(7):1715–1736.

Lütkephol, H. (2005). *New Introduction to Multiple Time Series Analysis.* Springer-Verlag: Berlin Heidelberg.

Mountford, A. and Uhlig, H. (2008). What are the effects of fiscal policy shocks? *NBER Working Papers*, 14551.

Oxley, L., Wilson, G., and Reale, M. (2009). Constructing structural VAR models with Conditional Independence Graphs. *Mathematics and Computers in Simulations*, 79:2916–2916.

Perotti, R. (2005). Estimating the effects of fiscal policy in oecd countries. *Proceedings of the Federal Bank of San Francisco.*

Perotti, R. (2007). In search of the transmission mechanism of fiscal policy. *NBER Macroeconomics Annual*, 22:169–226.

Ramey, V. and Shapiro, M. (1998). Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Confer. Series on Public Policy*, 48:145–194.

Ravn, M., Schmitt-Grohe, S., and Uribe, M. (2006). Deep habits. *Review of Economic Studies*, 73:195–218.

Reale, M. and Wilson, G. (2001). Identification of vector AR models with recursive structural errors using conditional independence graphs. *Statistical Methods and Applications*, 10(1):49–65.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419.

Wilson, G. and Reale, M. (2008). The sampling properties of conditional independence graphs for I (1) structural VAR models. *Journal of Time Series Analysis*, 29(5):802–810.

Appendix: Identification of SVAR models

Baseline:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -a_{21} & 1 & 0 \\ 0 & -a_{32} & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^t \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^t \end{bmatrix}}_{e_t} \tag{A1}$$

Consumption:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ 0 & -a_{42} & 0 & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^t \\ u_t^c \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^t \\ e_t^c \end{bmatrix}}_{e_t} \tag{A2}$$

Hours worked:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ 0 & -a_{42} & -a_{43} & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^h \\ u_t^t \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^h \\ e_t^t \end{bmatrix}}_{e_t} \tag{A3}$$

Real wage:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ -a_{41} & -a_{42} & 0 & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^t \\ u_t^w \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^t \\ e_t^w \end{bmatrix}}_{e_t} \tag{A4}$$

Residential investment:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ 0 & -a_{32} & 1 & 0 \\ 0 & -a_{42} & 0 & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^t \\ u_t^R \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^t \\ e_t^R \end{bmatrix}}_{e_t} \tag{A5}$$

Non-residential investment:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ -a_{31} & -a_{32} & 1 & 0 \\ 0 & -a_{42} & -a_{43} & 1 \end{bmatrix}}_{A_0} \underbrace{\begin{bmatrix} u_t^g \\ u_t^y \\ u_t^{NR} \\ u_t^t \end{bmatrix}}_{u_t} = \underbrace{\begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} e_t^g \\ e_t^y \\ e_t^{NR} \\ e_t^t \end{bmatrix}}_{e_t} \tag{A6}$$

Paper	Output	Consumption	Hours	Wage	Investment
Theoretical models:					
Baxter and King (1993)	+	-	+	-	+
Edelberg, Eichenbaum and Fisher (1999)	+	-	+	-	+ -*
Burnside, Eichenbaum and Fisher (2004)	+	-	+	-	+
Linneman (2006)	+	+	+	+	
Ravn, Schmitt-Grohe and Uribe (2006)	+	+	+	+	
Gali, Lopez-Salido and Valles (2007)	+	+	+	+	-
Empirical analyses:					
Ramey and Shapiro (1998)	+	-	+	-	
Edelberg, Eichenbaum and Fisher (1999)	+	-	+	-	+ -*
Blanchard and Perotti (2002)	+	+			-
Burnside, Eichenbaum and Fisher (2004)	+	-	+	-	+
Perotti (2005)	+	+			
Perotti (2007)	+	+			
Caldara and Kamps (2008)	+	+	+	+	
Mountford and Uhlig (2008)	+	+	+	-	-

* Edelberg, Eichenbaum and Fisher (1999) find that while the impact response of non-residential investment is positive, residential investment has instead a negative response.

Table 1: The sign of impact responses of key macroeconomic variables to a fiscal expansion according to theoretical and empirical studies.

	u_t^g	u_t^y	u_t^i
u_t^g	1.000		
u_t^y	0.248***	1.000	
u_t^i	0.016	0.395***	1.000

(a) baseline

	u_t^g	u_t^y	u_t^i	u_t^c
u_t^g	1.000			
u_t^y	0.143**	1.000		
u_t^i	0.026	0.347***	1.000	
u_t^c	0.064	0.638***	-0.104	1.000

(b) consumption

	u_t^g	u_t^y	u_t^i	u_t^h
u_t^g	1.000			
u_t^y	0.216***	1.000		
u_t^i	-0.000	0.170**	1.000	
u_t^h	0.016	0.558***	0.250***	1.000

(c) hours worked

	u_t^g	u_t^y	u_t^i	u_t^w
u_t^g	1.000			
u_t^y	0.257***	1.000		
u_t^i	0.030	0.384***	1.000	
u_t^w	-0.146**	0.150**	0.043	1.000

(d) real wage

	u_t^g	u_t^y	u_t^i	u_t^R
u_t^g	1.000			
u_t^y	0.305***	1.000		
u_t^i	0.015	0.291***	1.000	
u_t^R	-0.099	0.350***	0.041	1.000

(e) residential investment

	u_t^g	u_t^y	u_t^i	u_t^{NR}
u_t^g	1.000			
u_t^y	0.429***	1.000		
u_t^i	0.091	0.170**	1.000	
u_t^{NR}	-0.363***	0.656***	0.169**	1.000

(f) non-residential investment

Note: *,** and *** denote significance at 0.10, 0.05 and 0.01 levels, respectively. The corresponding threshold values for the baseline model are 0.1189, 0.1408 and 0.1840, respectively. For all the other models, they are 0.1204, 0.1426 and 0.11864, respectively.

Table 2: Estimated partial correlations of the series innovations.

Model	AIC	HIC	SIC
A	506.00	530.16	565.73
C	507.04	531.20	566.76
D	571.64	551.80	587.37

Table 3: Information criteria associated to the feasible DAGs of the baseline model.

	ε_t^g	ε_t^y	ε_t^i
ε_t^g	1.000		
ε_t^y	0.000	1.000	
ε_t^i	0.066	-0.008	1.000

(a) baseline

	ε_t^g	ε_t^y	ε_t^i	ε_t^c
ε_t^g	1.000			
ε_t^y	0.000	1.000		
ε_t^i	0.018	-0.005	1.000	
ε_t^c	0.060	-0.013	-0.093	

(b) consumption

	ε_t^g	ε_t^y	ε_t^i	ε_t^h
ε_t^g	1.000			
ε_t^y	0.000	1.000		
ε_t^i	0.000	0.000	1.000	
ε_t^h	0.015	-0.004	0.000	1.000

(c) hours worked

	ε_t^g	ε_t^y	ε_t^i	ε_t^w
ε_t^g	1.000			
ε_t^y	0.000	1.000		
ε_t^i	0.024	-0.007	1.000	
ε_t^w	-0.000	0.000	0.044	1.000

(d) real wage

	ε_t^g	ε_t^y	ε_t^i	ε_t^R
ε_t^g	1.000			
ε_t^y	0.000	1.000		
ε_t^i	0.011	-0.035	1.000	
ε_t^R	0.094	0.031	0.040	1.000

(e) residential investment

	ε_t^g	ε_t^y	ε_t^i	ε_t^{NR}
ε_t^g	1.000			
ε_t^y	0.000	1.000		
ε_t^i	0.081	-0.026	1.000	
ε_t^{NR}	-0.000	0.000	0.032	1.000

(f) non-residential investment

Note: The two-standard-error band for a sample size of 204 is ± 0.1400.

Table 4: Correlations between residuals of the DAGs fitted to the VAR(4) estimated innovations.

Models	X_t	$-a_{21}$	$-a_{31}$	$-a_{32}$	$-a_{41}$	$-a_{42}$	$-a_{43}$	b_{11}	b_{22}	b_{33}	b_{44}	m	LR-$\chi^2(m)$
Baseline	$gy\,t$	-0.22***	0	-1.90***	-	-	-	1.10***	0.82***	3.42***	-	1	0.27
		(0.05)	-	(0.28)	-	-	-	(0.05)	(0.04)	(0.17)	-		[0.87]
Consumption	$gy\,t\,c$	-0.20***	0	-1.72***	0	-0.53***	0	1.11***	0.80***	3.34***	0.49***	3	2.29
		(0.05)	-	(0.28)	-	(0.04)	-	(0.05)	(0.04)	(0.16)	(0.02)		[0.51]
Hours worked	$gy\,h\,t$	-0.23***	0	-0.004***	0	-0.95***	-193.75***	1.10***	0.81***	0.004***	3.26***	2	0.07
		(0.05)	-	(0.0004)	-	(0.35)	(51.24)	(0.05)	(0.04)	(0.0002)	(0.16)		[0.97]
Real wage	$gy\,t\,w$	-0.22***	0	-1.91***	0.001	-0.001**	0	1.09***	0.82***	3.43***	0.01***	2	0.14
		(0.05)	-	(0.28)	(00004)	(0.0005)	-	(0.05)	(0.04)	(0.17)	(0.0003)		[0.93]
Residential investment	$gy\,t\,R$	-0.23***	0	-1.65***	0	-1.66***	0	1.10***	0.76***	3.35***	3.41***	3	2.59
		(0.05)	-	(0.29)	-	(0.30)	-	(0.05)	(0.04)	(0.17)	(0.17)		[0.46]
Non-residential investment	$gy\,NR\,t$	-0.23***	1.35***	-4.78***	-	-1.41***	-0.11**	1.10***	0.78***	3.69***	3.32***	1	1.44
		(0.05	(0.25)	(0.33)	0	(0.38)	(0.06)	(0.05)	(0.04)	(0.18)	(0.16)		[0.23]

Note: Estimation method: maximum likelihood. Standard errors in parentheses. *,** and *** denote significance at a 0.10, 0.05 and 0.01 level respectively. m = number of over-identifying restrictions. LR = likelihood-ratio χ^2 statistics. P-values in square brackets.

Table 5: Structural factorisations.

		Excluded period				
	full sample	1955:1-1964:4	1965:1-1974:4	1975:1-1984:4	1985:1-1994:4	1995:1-2006:4
Government spending shock						
Government spending	1.15*	1.24*	1.00*	1.11*	1.03*	1.15*
	(4)	(4)	(1)	(4)	(2)	(4)
Output	1.75*	1.42*	1.59*	1.76*	2.24*	1.31*
	(8)	(3)	(5)	(3)	(4)	(10)
Tax	0.51*	0.58*	-0.49	0.54*	1.20*	1.05*
	(6)	(20)	(2)	(6)	(10)	(11)
Consumption	0.87*	0.91*	0.82*	1.15*	1.63*	0.71*
	(10)	(20)	(1)	(2)	(4)	(10)
Hours worked	0.15*	0.12*	-0.08	0.20*	0.34*	0.29*
	(6)	(3)	(2)	(3)	(4)	(11)
Real wage	0.24*	0.23*	0.73*	0.37*	0.34*	0.31*
	(15)	(8)	(1)	(4)	(8)	(19)
Residential investment	0.20*	0.22*	0.25*	0.29*	0.25*	0.26*
	(5)	(12)	(5)	(2)	(12)	(8)
Non-residential investment	-0.73*	-0.75*	0.70*	0.50	0.18	-0.41
	(2)	(11)	(7)	(3)	(8)	(2)
Tax shock						
Government spending	-0.07	-0.09	0.18*	0.15	-0.11*	-0.12*
	(6)	(7)	(19)	(18)	(26)	(15)
Output	0.17	-0.24	0.45*	-0.39	-1.16*	-0.63
	(3)	(13)	(3)	(11)	(11)	(8)
Tax	1.00*	1.00*	1.00*	1.15*	1.00*	1.00*
	(1)	(1)	(1)	(3)	(1)	(1)
Consumption	-0.11*	-0.13	0.24	-0.61*	-0.73*	-0.37*
	(2)	(8)	(3)	(12)	(10)	(9)
Hours worked	-0.04	-0.06	-0.05	0.09	-0.09	-0.09
	(15)	(16)	(15)	(5)	(12)	(9)
Real wage	0.03	0.07*	0.20*	0.12*	0.02	-0.06*
	(10)	(10)	(17)	(20)	(4)	(10)
Residential investment	-0.08	-0.12	-0.18*	-0.12*	-0.30*	-0.21*
	(5)	(6)	(8)	(8)	(7)	(6)
Non-residential investment	0.07*	0.37*	0.36*	0.35*	0.36*	0.19
	(3)	(3)	(3)	(3)	(3)	(3)

Note: Quarters in parentheses. * denotes significance at a 0.10 level.

Table 6: Subsample stability: peak responses.

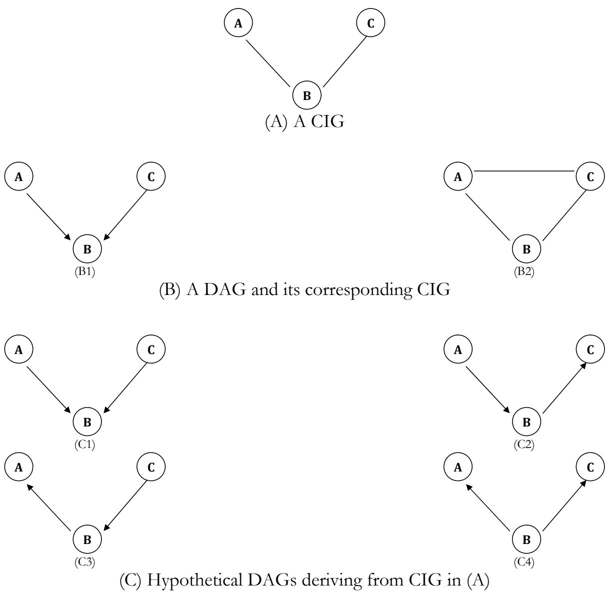


Figure 1: Conditional independence graphs and directed acyclic graphs.

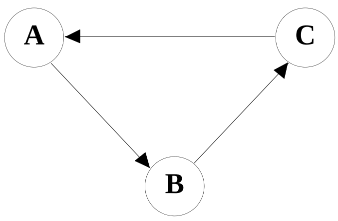
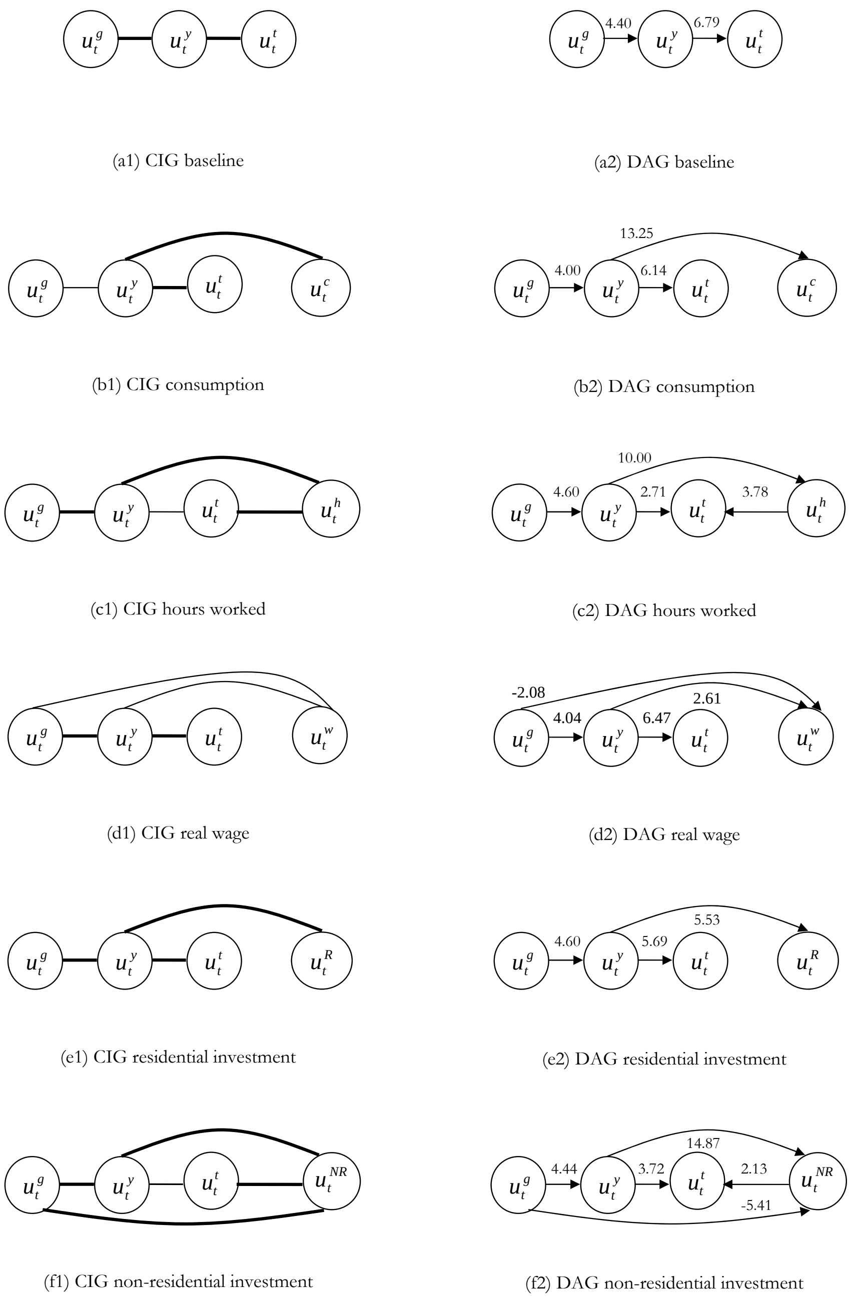


Figure 2: Directed cyclic graph.



(a1) CIG baseline (a2) DAG baseline

(b1) CIG consumption (b2) DAG consumption

(c1) CIG hours worked (c2) DAG hours worked

(d1) CIG real wage (d2) DAG real wage

(e1) CIG residential investment (e2) DAG residential investment

(f1) CIG non-residential investment (f2) DAG non-residential investment

Note: In the CIGs (left column), the strengths of the links are indicated by significance at the 0.10 level (dashed line), 0.05 level (thin line), 0.01 level (bold line). In the DAGs (right column), t-statistics of the estimated regression coefficients are shown adjacent to the directed links.

Figure 3: Sample CIGs and estimated DAGs fitted to VAR(4) residuals.

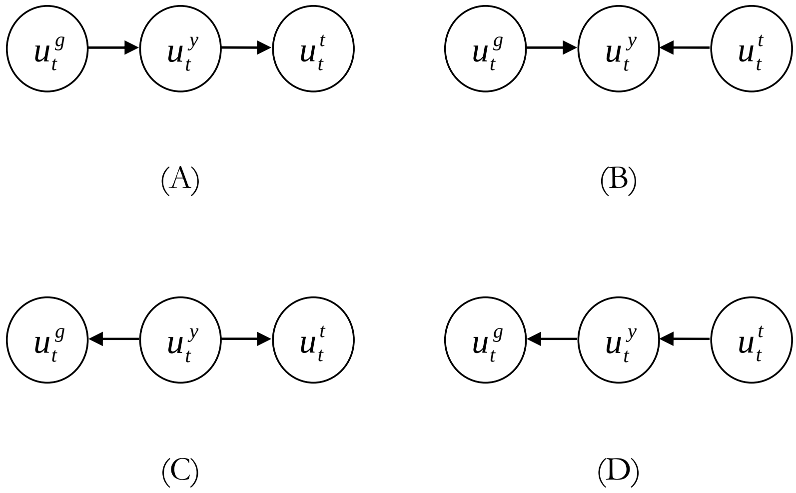
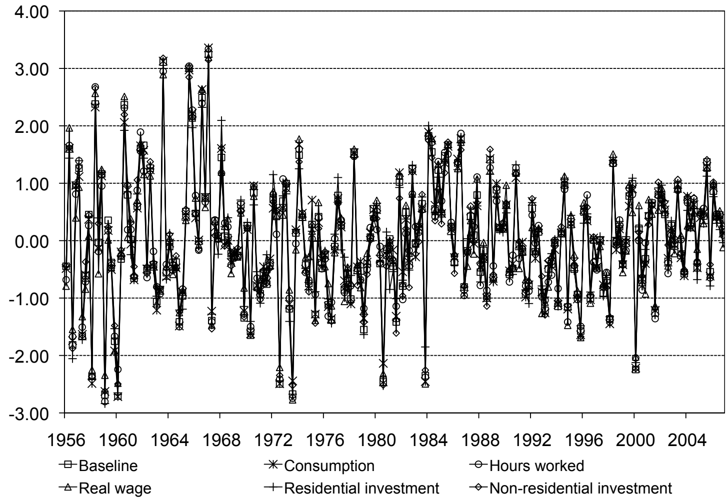
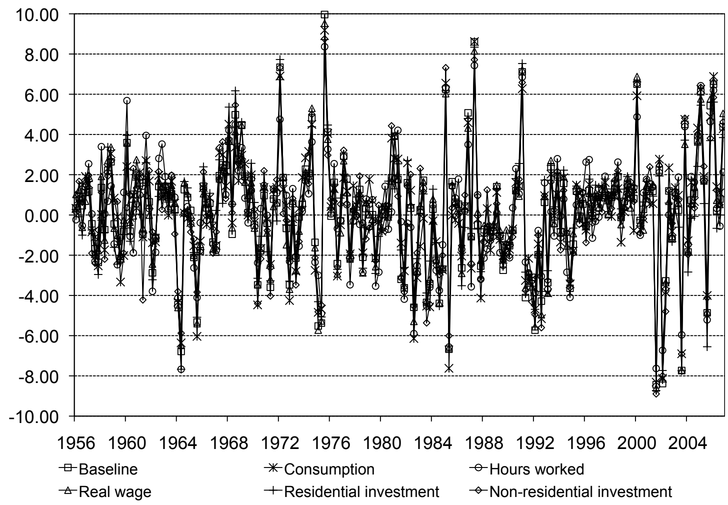


Figure 4: All possible DAGs deriving from the CIG of the baseline model.

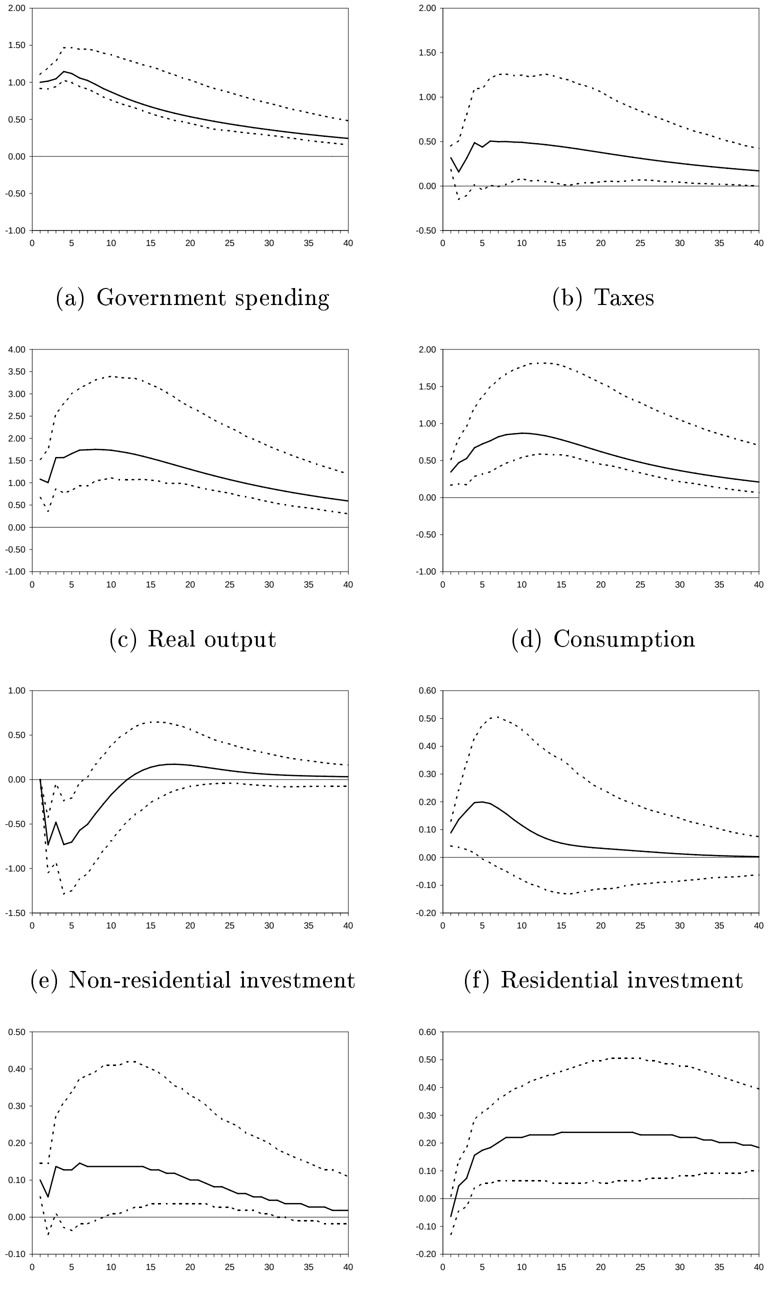


(a) Government spending shocks.



(b) Tax shocks.

Figure 5: Identified fiscal policy shocks across models.



(a) Government spending

(b) Taxes

(c) Real output

(d) Consumption

(e) Non-residential investment

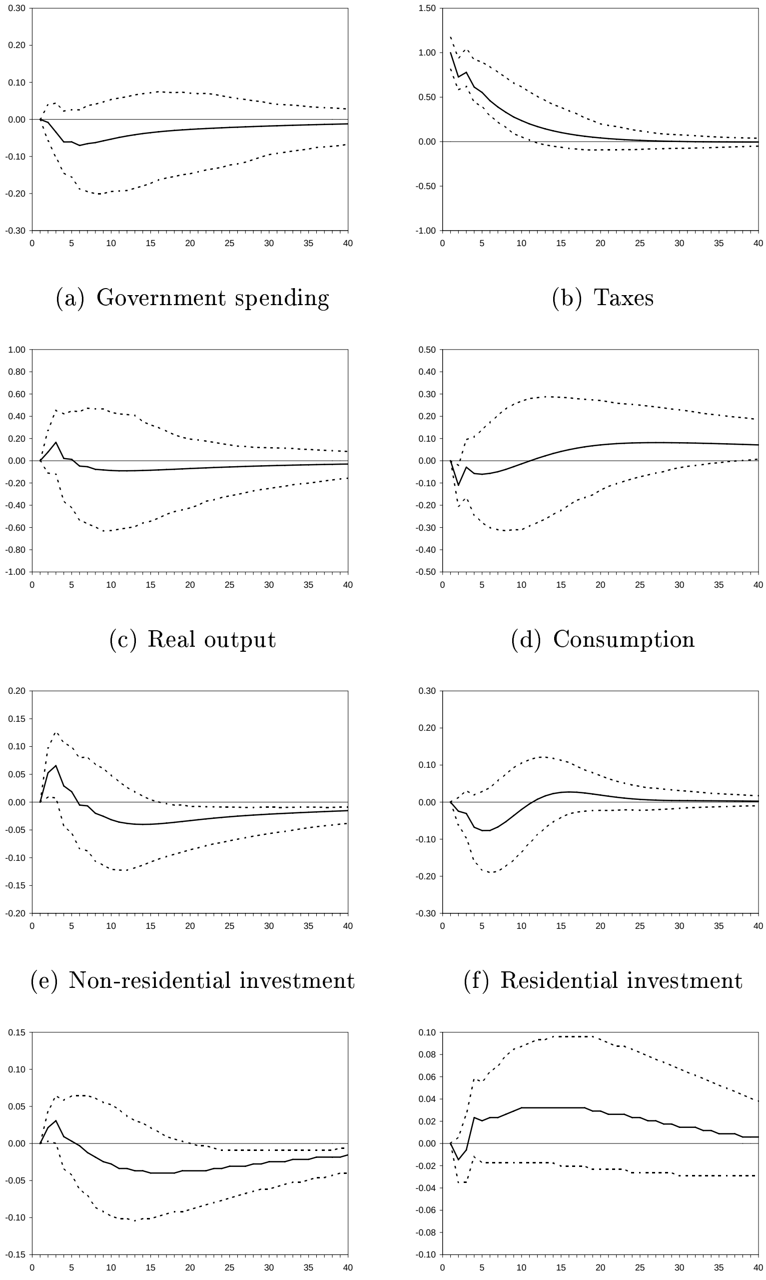
(f) Residential investment

(g) Hours worked

(h) Real wages

Note: Dashed lines represent 90% confidence intervals computed according to Hall's (1992) algorithm with 2000 bootstrap replications. Responses are shown for a 40-quarter horizon. The impulse responses of government spending, taxes and real output are computed on the basis of the baseline 3-variable SVAR. The impulse responses of the remaining variables are obtained from 4-variable models obtained by adding one variable at a time to the baseline model. Impulse response of real output and its components are rescaled to represent the dollar change of the variables to a shock to government spending of size one dollar. The impulse responses of hours worked and real wages are rescaled to represent the percentage change subsequent to a government spending shock of size one percent.

Figure 6: Impulse responses to a government spending shock.



(a) Government spending

(b) Taxes

(c) Real output

(d) Consumption

(e) Non-residential investment

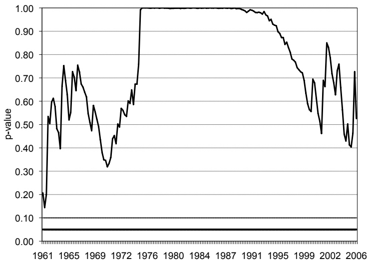
(f) Residential investment

(g) Hours worked

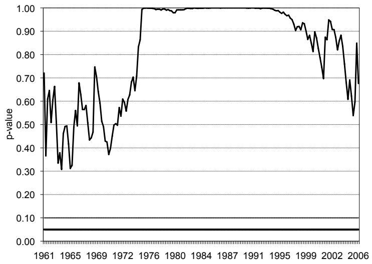
(h) Real wages

Note: Dashed lines represent 90% confidence intervals computed according to Hall's (1992) algorithm with 2000 bootstrap replications. Responses are shown for a 40-quarter horizon. The impulse responses of government spending, taxes and real output are computed on the basis of the baseline 3-variable SVAR. The impulse responses of the remaining variables are obtained from 4-variable models obtained by adding one variable at a time to the baseline model. Impulse response of real output and its components are rescaled to represent the dollar change of the variables to a shock to tax revenues of size one dollar. The impulse responses of hours worked and real wages are rescaled to represent the percentage change subsequent to a tax revenue shock of size one percent.

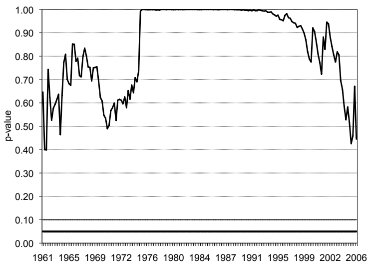
Figure 7: Impulse responses to a tax shock.



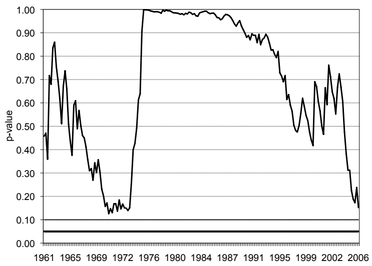
(a) Baseline



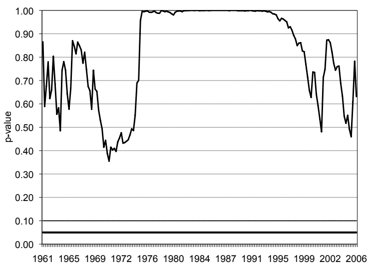
(b) Consumption



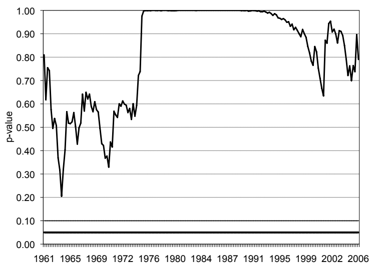
(c) Hours worked



(d) Real wage



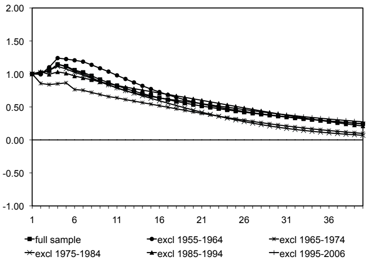
(e) Non-residential investment



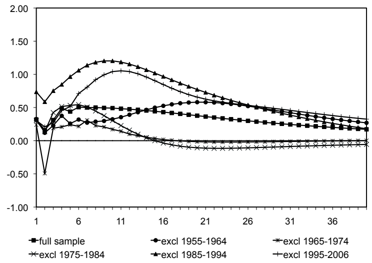
(f) Residential investment

Note: Bold horizontal lines represent the 0.05 significance level. Thin horizonal lines represent the 0.10 significance level. Chow forecast test recursively run at every quarter from 1961:3 to 2006:3. Null hypothesis: parameter constancy. P-values computed with 2000 bootstrap replications.

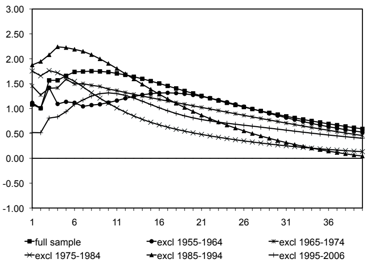
Figure 8: Chow forecast test (recursive p-values).



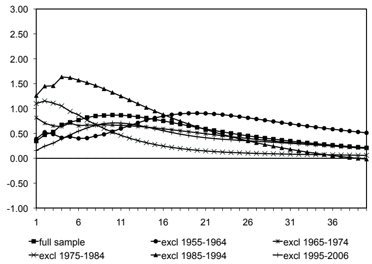
(a) Government spending



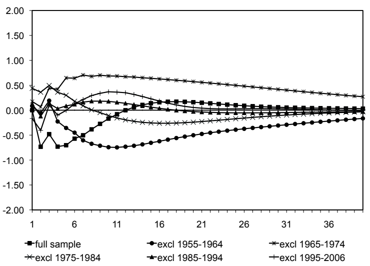
(b) Taxes



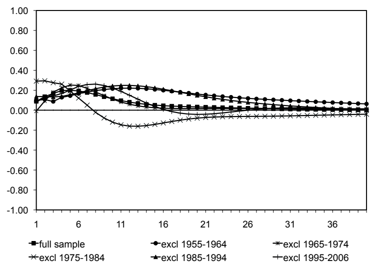
(c) Real output



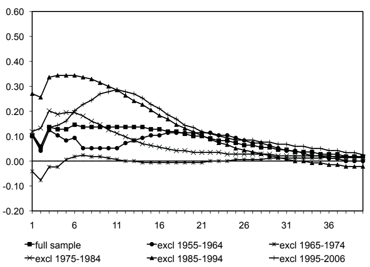
(d) Consumption



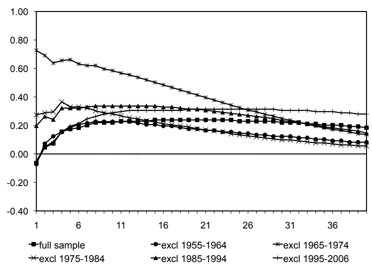
(e) Non-residential investment



(f) Residential investment

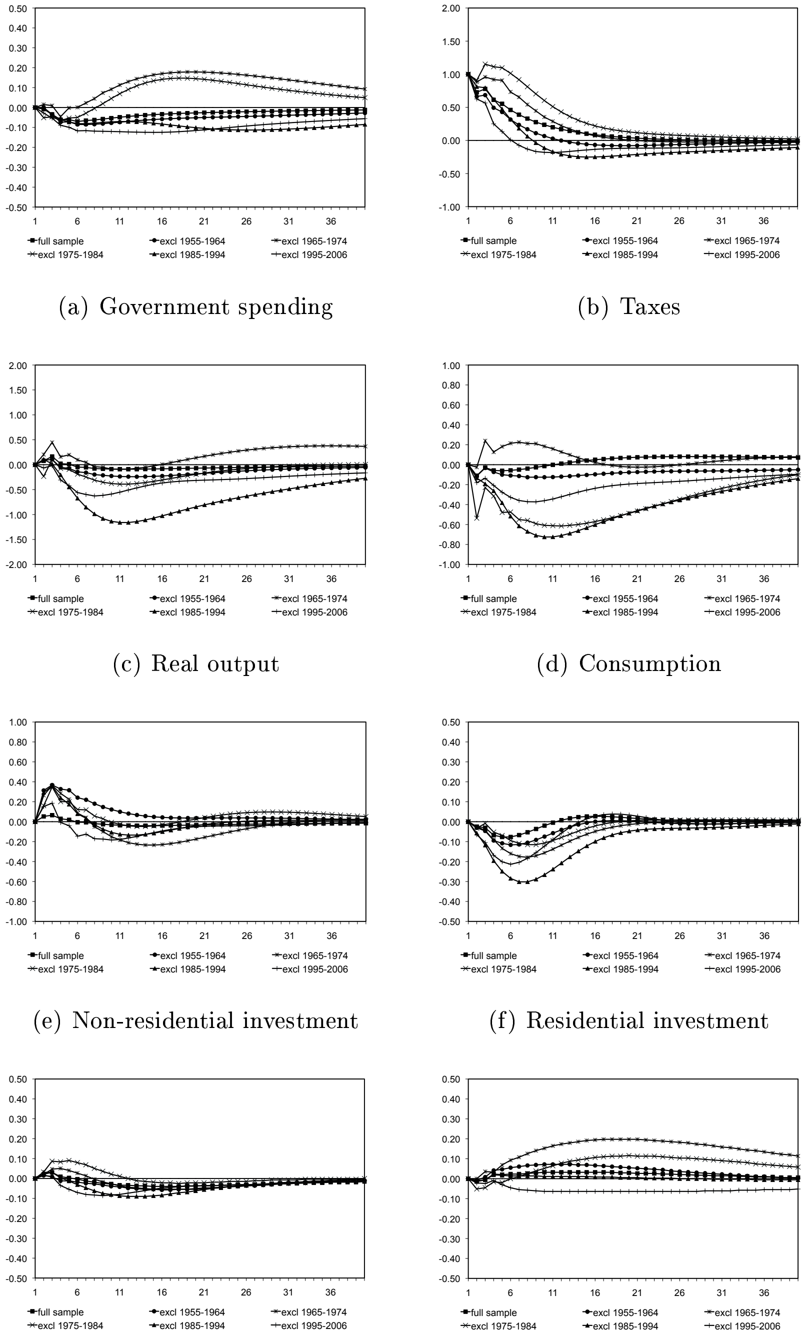


(g) Hours worked



(h) Real wages

Figure 9: Subsample stability: government spending shock.



(a) Government spending

(b) Taxes

(c) Real output

(d) Consumption

(e) Non-residential investment

(f) Residential investment

(g) Hours worked

(h) Real wages

Figure 10: Subsample stability: tax shock.