ufficio parlamentare di bilancio

upB. ufficio parlamentare di bilancio

WORKING PAPER

Fiscal Multipliers in Abnormal Times: the Case of a Model of the Italian Economy

by Sergio De Nardis and Carmine Pappalardo

Working Paper no. 1 / *December 2018*

The Working Papers series comprises technical and institutional studies conducted by the PBO, with the collaboration of independent experts where appropriate. The Working Papers seek to explore in greater detail specific issues relevant to the institutional work of the PBO.

The contents of the Working Papers represent the views of their authors and do not necessarily reflect those of the PBO or the institutions with which the independent experts are affiliated.

The electronic version can be downloaded from: www.upbilancio.it

Parliamentary Budget Office
Via del Seminario, 76
00186 Rome, Italy
segreteria@upbilancio.it

Fiscal Multipliers in Abnormal Times: the Case of a Model of the Italian Economy

by Sergio de Nardis, Carmine Pappalardo[1]

Abstract

In this paper we provide estimates of the fiscal multipliers for the Italian economy in the crisis period (2008-2014). Based on a traditional structural macro-econometric model, we find suggestive evidence of an increase in the size of such multipliers in the latter part of the sample, involving the exceptional period of double-dip recession (2008-2014). How to get from these indications to a more precise inference of crisis-specific multipliers is an unresolvable problem within standard model estimations because the timespan is too short to obtain any reliable and efficient inference of the crisis multipliers. We circumvent this problem by first correcting the model for any instability in the structural parameters. Then, the indirect inference of fiscal multipliers for the crisis period (2008-2014) is obtained on the basis of an (inverse) variance-based weighting scheme of the sub-periods fiscal multipliers. We show that, despite the higher statistical uncertainty, the magnitude of the estimated multipliers of the crisis period is significantly larger than pre-crisis estimates, both on the expenditure and the revenue sides, and the findings are robust to sensitivity checks. Results from the simulation of the 2011-14 fiscal consolidation plan show that appropriate consideration of the crisis-specific multipliers considerably reduces the forecast error with respect to the projection obtained on the grounds of the standard model-based multipliers.

[1] Sergio De Nardis, Parliamentary Budget Office and Luiss School of European Political Economy; Carmine Pappalardo, Parliamentary Budget Office, Rome, Italy. We are grateful to Roberto Basile, Davide Fantino, Roberto Golinelli, Libero Monteforte, Guseppe Pisauro, Pietro Rizza and Stefano Siviero for their comments and very helpful suggestions to improve the paper. All remaining errors are ours. The views expressed in this work are our own and do not necessarily reflect those of Parliamentary Budget Office.

Non technical summary

In this paper we provide estimates of the fiscal multipliers for the Italian economy in the crisis period (2008-2014), based on a traditional structural macro-econometric model (MeMo-It model).

We start by performing a sub-period estimation of the model in 1970-2007, that is for the years preceding the crisis. We find that the fiscal multipliers obtained from this model estimation are generally smaller than the fiscal multipliers of the model estimated for the entire sample period 1970-2014, which includes the crisis period (2008-2014). We interpret this evidence as suggestive of an increase in the size of fiscal multipliers in the latter part of the sample, involving the crisis period.

Yet, how we go from these suggestive indications to a more precise inference of crisis-specific multipliers is an unresolvable problem within standard model estimations. This is because: *a)* model-based estimates of fiscal multipliers in these models are, by construction, independent of the state of the economy (they are the same in recessions and expansions); *b)* the timespan of the crisis is too short to make any accurate and efficient direct estimate of the multipliers in those years feasible.

We show however that even such models convey, through the inspection of stability of structural parameters and sub-period estimations, relevant information about changes in fiscal multipliers in the crisis period. The methodology we adopt consists of three steps:

a) Correction of the model for any instability of structural parameters (in both cointegration relations and error correction equations);
b) Re-estimation of the fiscal multipliers of the instability-amended model both for the whole sample period (1970-2014) and the pre-crisis years (1970-2007);
c) Indirect estimation of fiscal multipliers for the crisis period (2008-2014) on the basis of the construction of (inverse) variance-based weighting schemes of the sub-periods fiscal multipliers.

Although uncertainty is inevitably large in this kind of estimation, it is nonetheless possible to show that the size of the multipliers of the crisis period (2008-2014) is in general significantly larger than pre-crisis estimates (1970-2007), both on the expenditure and the revenue sides. According to the baseline estimates, the size gets quite large in the crisis period for multipliers associated with government expenditure (they rise above 2.5 for investment expenditure and above 1 for consumption). Significant increases in crisis-period multipliers (which remain below 1 however) are also detected for expenditures on social transfers, direct taxation (both household and corporate income), social security contributions and taxation of consumption and (regional) economic activities. Differences are still detected even after various sensitivity tests, confirming, for the Italian case, the findings of the empirical literature about the increase in the size of fiscal multipliers in the crisis period compared with normal times.

As an application, we adopt the estimated (baseline) period-specific multipliers for the multi-year fiscal consolidation plan implemented, in successive instalments, by Italy in the second half of 2011. Particularly, we decompose the forecast error with respect to a projection obtained with the MeMo-It model based on its standard multipliers in three components of errors respectively due to: *a)* exogenous variables; *b)* fiscal multipliers; and *c)* other factors. Simulation results show the large role of fiscal multipliers in the forecast error concerning the depth and length of the subsequent recession. Correcting the underestimation of fiscal multipliers makes the forecast error almost vanish in 2012 and decline by 70 percent in 2013. Robustness checks, performed on the assumption of positive values of the covariance between the same multipliers in different periods, substantially confirm the baseline result of a noticeable reduction in the forecast error.

1 Introduction

The Great Recession and its European extension with the sovereign debt crisis and the fiscal adjustments undertaken in several Eurozone countries have rekindled attention on the issue of the effectiveness of fiscal policy. As Fatàs and Mihov (2012) point out, the pre-crisis consensus view of both academic researchers and policymakers focused almost exclusively on monetary policy as a stabilization tool, assuming that fiscal policy was not good at doing the job. More specifically, it was thought that automatic stabilizers could generally provide a degree of stabilization. When this was not enough, discretionary fiscal policy had to be put aside in favor of monetary policy because of the lags in the decision-making and implementation of fiscal measures and the distortions induced by political interference. According to Blanchard at al. (2010), fiscal policy took a backseat to monetary policy for several other reasons. First, there was wide skepticism about the effects of fiscal policy, also on the grounds of arguments related to Ricardian equivalence. Moreover, in the pre-crisis period of "great moderation", monetary policy had proved successful in maintaining a stable output gap and hence there was little need for another stabilization instrument. Finally, in the advanced countries affected by aging populations and problems of the long-term sustainability of the public finances, the priority of fiscal policy had to be to stabilize and possibly reduce high debt levels. Also based on these views, a large majority of analysts and policy-makers anticipated somewhat limited and short-lived recessionary effects from the fiscal consolidation plans that, absent any sort of backstop, several Eurozone economies had to frontload in response to the burst of financial panic and the fall of confidence in the single currency.

Yet, the general tenet about the conduct of economic policy was dependent on the realization of conditions that hold in normal times, but which failed to materialize in the exceptional circumstances of the crisis. Consolidation plans were adopted at a time of considerable slack in economic activity (negative output gaps), when European economies had not yet recovered from the earlier severe downswing. Fiscal adjustments were therefore implemented in recessionary time. In this situation, monetary policy became rapidly impotent in regulating the business cycle, as central banks (and particularly the ECB) could not cut interest rates much below the zero lower bound to counteract the negative impact of fiscal consolidations. Quite the contrary, the deflationary impulses produced by austerity plans widened the gap between the real interest rates and their (falling) equilibrium levels. Furthermore, the impairment of financial systems (the credit crunch) significantly increased the share of financially constrained agents, which reacted more strongly to the squeeze on their current income than in normal times. On the top of that, fiscal adjustments were implemented simultaneously across Europe, even in countries that were not facing confidence crises. The lack of fiscal policy coordination notably intensified the negative spillover effects in an area characterized by strong trade links. In their influential work, Blanchard and Leigh (2013) show that, due to the underestimation of fiscal multipliers based on pre-crisis developments, GDP forecast errors (made by IMF and other international organizations) were significant and widespread in the advanced countries and the euro area. In other words, fiscal multipliers in the crisis period were substantially higher than those assumed by forecasters on the basis of estimates performed in normal times. A similar finding was more recently confirmed by Górnicka et al. (2018), who found that the fiscal multipliers applied by the European Commission increased over time during the crisis period.

In this paper we address these issues in the case of the Italian economy. To our knowledge the question of rising fiscal multipliers in the crisis years has indeed been much discussed in the Italian policy debate, but has been scarcely dealt with by the main workhorse models used by institutions in their forecasting activity. We take on this issue working with the Istat macro-econometric model (MeMo-It), currently used by the Parliamentary Budget Office (PBO) for its own forecasting activity.[2] This is a traditional structural econometric model and as such it does not allow, by construction, the estimation of response functions to the impulse of any fiscal policy variable that are dependent on the state of the economy. Nonetheless, we show that even such models can convey, through the inspection of parameter stability and sub-period estimations, relevant information about possible changes in fiscal multipliers in the crisis period. We make use of this information by adapting older statistical methodologies (pioneered by Chow, 1960) to the construction of (inverse) variance-based weighting schemes of sub-period estimates of fiscal multipliers to infer their (implicit) dimension in a period (the crisis years) for which a direct estimation is not feasible. The approach we follow is close to spirit of Blanchard-Leigh, in the sense that, as in their work, it can provide ex-post evidence of crisis-specific fiscal multipliers by exploiting the information generated by the analysis of the period-specific performance changes of state-independent models.

Although uncertainty is inevitably large in this kind of estimation, it is nonetheless possible to show a generally large and statistically significant increase in the dimension of fiscal multipliers implicit in the crisis period (2008-2014) compared with pre-crisis model estimates. According to our baseline estimates for the crisis period, size gets quite large for multipliers associated to government expenditures (they increase to above 2.5 for investments and above 1 for consumption). Significant increases in crisis-period multipliers (remaining below 1 however) are also detected for expenditure on social transfers, direct taxation (both household and corporate income), social security contributions and taxation of consumption and (regional) economic activities. Making use of these estimates, we also show that the application of the crisis-specific fiscal multipliers to the multi-year consolidation plan adopted, in successive instalments, by Italy in the second half of 2011 (the Berlusconi-Monti fiscal adjustment) considerably reduces the forecast error with respect to a projection obtained with the MeMo-It model based on its standard multipliers. Correction of the underestimation of fiscal multipliers makes the forecast error about the depth of the post-austerity recession almost vanish in 2012 and decrease by 70 percent in 2013.

[2] The PBO uses that model under the terms of a framework agreement signed with the National Statistical Institute (Istat); http://en.upbilancio.it/wp-content/uploads/2015/04/Accordo-tra-Upb-e-Istat.pdf

The paper is organized as follows. In Section 2 a survey illustrates the state of discussion and the evidence in the literature about the size of fiscal multipliers. In Section 3, the main features of the MeMo-It model are briefly described, providing an overview of its fiscal multipliers and preliminary evidence pointing to possible changes in the last few (crisis) years. Stability tests of structural parameters are performed in Section 4, where fiscal multipliers are re-estimated, after controlling for structural breaks, considering both the whole sample size and the pre-crisis period and checking for statistical significance of the detected differences. Section 5 is devoted to describing the methodology adopted to infer the fiscal multipliers over the crisis period (2008-2014), showing the baseline estimates and performing several sensitivity tests of the base-results. In Section 6 we perform a predictive validation exercise for these estimates, applying the (baseline) crisis-specific multipliers to the 2011 fiscal consolidation plan and showing the magnitude of correction of the forecast error allowed by these estimates. Section 7 concludes.

2 Review of the literature on the size of fiscal multipliers

Fiscal multipliers are summary measures of the output response to exogenous (discretionary) impulses of fiscal policy. They are defined as the ratio of the variation in output to the discretionary change in the relevant fiscal variable with respect to a baseline scenario (Spilimbergo et al. 2009). Given this definition, fiscal multipliers are a function of the structural parameters and the policy-reaction parameters of the underlying model (Chinn 2013).

In theory, fiscal multipliers can be of any dimension depending on the assumptions the researcher adopts on the degree of flexibility of wages and prices, the optimizing behavior of agents (Ricardian equivalence; Barro 1974 and 1989) and the degree of agents' heterogeneity in facing financing constraints. Given such theoretical indeterminacy, the issue of the magnitude of fiscal multipliers boils down to an empirical matter. The literature of applied studies is quite voluminous and has been flourishing in the last few years, when the explosion of the financial crisis and its prolongation in Europe seemed to replicate the conditions of the Great Depression.

There is broad empirical consensus that the size of fiscal multipliers depends on both the structural characteristics of countries and the state of their business cycle (Batini et al. 2014a; Mineshima et al. 2014).

As for country characteristics, fiscal multipliers are larger in economies with a smaller propensity to import (that is, in large economies and in those relatively closed to trade). This is because the spending leakage on foreign goods due to imports rises with income, reducing in the multiplication process the increase in economic activity induced by an initial fiscal stimulus (Ilzetzki et al. 2013). Moreover, economies with a fixed exchange rate regime tend to have a larger multiplier than those operating under flexible exchange rates. This fact is related to the Mundell-Fleming argument according to which, under fixed exchange rates, the monetary authority has to expand the money supply following a fiscal stimulus just to prevent the exchange rate from appreciating (Ilzetzki et al. 2009). Similar results are also obtained controlling for other factors (public debt levels, condition in the financial system), although through different transmission mechanisms (Corsetti et al. 2012). For reasons akin to those detected for the case of fixed exchange rate regimes, economies that are member of a single currency area also tend to have larger fiscal multipliers than stand-alone countries (Nakamura and Steinsson 2014).

Besides the international links, the domestic structural features of economies play an equally important role in differentiating the magnitude of fiscal multipliers. More specifically, countries with more rigid labor markets can have larger fiscal multipliers to the extent that such rigidities translate in smaller wage responses when a fiscal policy shock hits demand (Cole and Ohanian 2004; Auerbach and Gorodnichenko 2012a). The size of the automatic stabilizers also affects the impact of fiscal policy, since small stabilizers imply limited automatic responses of tax revenues and transfer expenditures to income increases induced by a discretionary fiscal stimulus, amplifying the dimension of fiscal multipliers (Dolls et al. 2012). Finally, low-debt countries have generally larger fiscal multipliers than high-debt countries, as a fiscal stimulus performed by the latter may have negative credibility and confidence effects that impact private demand. According to some studies, fiscal multipliers substantially are lower in high-debt countries (Ilzetzki et al. 2013; Corsetti et al. 2012; Kirchner at al. 2010).

Several recent empirical studies have shown that fiscal multipliers are not only space-dependent, changing according to the structural characteristics of countries, but are also time-dependent, changing within the same economy depending on the state of the business cycle (see e.g. Auerbach and Gorodnichenko 2012a; Baum et al. 2012; Karras 2014; Fazzari et al. 2015; Jorda and Taylor 2016; Boitani and Perdichizzi 2018; Arin at al. 2018). The dimension of fiscal multipliers is generally larger in recession than in expansion. A fiscal stimulus implemented in an upturn tends to be less effective because, as the economy reaches full employment, more government purchases can crowd out private demand, leaving the level of activity unchanged but at a higher level of wages and prices. Conversely, a fiscal stimulus during a downturn, when there is spare production capacity, can crowd in private demand, drawing into use idle productive inputs without causing inflation. Analogously, a fiscal consolidation is more costly in terms of output loss in a downturn than in an upturn, because during a recession there is a higher proportion of credit-constrained agents who adopt hand-to-mouth behavior (who spend their current income), being less able to borrow to smooth expenditure over their lifetime. The high impact of government expenditure in recessions may even offset the impact of the smaller multipliers characterizing high-debt countries, to the extent that positive expenditure shocks in these economies do not lead to higher debt-to-GDP ratios (Boitani and Perdichizzi 2018). Given these findings, a major consequence is that estimates of multipliers based on an average of periods, including both upturns and downturns of economic activity along the normal fluctuations of business cycle, may be less informative, particularly in deep recessions (Batini et al. 2014b).

The latter remark has become particularly relevant in the case of the recent financial crisis, when depression-like conditions gave rise to the exceptional circumstances epitomized by the Great Depression and thoroughly analyzed in Neo-Keynesian models (Woodford 2011). In particular, the emergence in several economies of the binding zero lower bound in nominal interest rates prevented central banks from cutting interest rates to offset the negative short-term impact of fiscal consolidations on output. Moreover, seriously impaired financial systems implied that, in a situation of falling output, consumption depended more on current than future income and investment was affected more by current than future profits. In addition, fiscal consolidation was pursued, particularly in Europe, in a situation of considerable slack in economies that were just emerging from the financial crisis and still had negative output gaps. All these conditions led to large fiscal multipliers, substantially larger than those assumed (explicitly or implicitly) by the forecasters at the time. The consequences were large forecast errors, with a substantial underestimation of the output fall caused by the fiscal consolidations adopted in the main economies (Blanchard and Leigh 2013). The finding that the fiscal multiplier is larger at the zero lower bound than in normal times and (for government spending) well above 1 is generally common to all the empirical testing inspired by the New-Keynesian DSGE models (Cogan et al. 2010; Christiano et al. 2011; Coenen at al. 2012). It is also shared by the empirical studies that do not find confirmation of a state-dependency of fiscal multipliers over the ups and downs of the business cycle (Ramey and Zubairy 2018).

Besides space- and state-dependency, fiscal multipliers exhibit significant heterogeneity across the different categories of public intervention, so that analyses of the effects of a fiscal stimulus/consolidation based only on considering the change in the overall government budget may be quite misleading. The composition of fiscal interventions is crucial for a correct assessment of the effect of a budget consolidation/expansion. In general, there is broad consensus on the view that, after controlling for the response of monetary policy and exchange rates (Perotti 2013; Guajardo et al. 2014), multipliers associated with government spending (purchases of goods and services and public investment) are larger, in the short run, than those associated with tax revenues (Blanchard and Perotti 2002; Mountford and Uhlig 2009). Moreover, spending multipliers tend to rise during downturns more than those associated with tax revenues, leading to an increase, in a recession, in the difference between the effects associated with the two different kinds of intervention (Gechert at al. 2016).

Expectations, confidence and agents' forecasts of the future supply-side effects of fiscal measures may also play a role in determining the size of multipliers, although the findings in the empirical literature are not unambiguous. According to one position, consideration of these factors leads to output impacts of fiscal consolidations that are the opposite of those envisaged by the majority of the consensus views. Fully anticipated and credible consolidation programs implemented through expenditure cuts have, independently of the state of the cycle, smaller recessionary impacts than those implemented through tax increases, thanks to the modification of expectations, which would favor an expansion in the former case, and a contraction in the latter (Alesina et al. 2015). This empirical finding is confirmed when one considers the multi-year nature of fiscal adjustments, affecting the planning of investors and consumers, and the interdependence of the government decisions about spending cuts and tax increases, although it does not seem to hold up when the zero lower bound of the policy interest rate is also considered (Alesina et al. 2018). Different conclusions are reached by another approach that hinges on agents' expectations, under which in a downturn it is the credible announcement of a future expansion of government spending that induces a higher fiscal multiplier, with the corollary that it is the agents' fiscal foresight about the persistence of the government spending increase over the entire recession that gives boost to the economy (Figueres 2015).

Finally, estimates of fiscal multipliers can be obtained in several ways, with the empirical approaches often associated with some particular theoretical framework. In general, three main estimation methodologies can be singled out. The first is based on traditional structural econometric models, a category of multi-equation macro-econometric models to which MeMo-It belongs, where fiscal multipliers are measured in terms of the response of output to shocks to the exogenous fiscal variables in the equations of the model. An alternative approach regards the vector auto-regression framework (VAR), where there are no exogenous variables and multipliers are estimated as the output response to the error term, so they measure the reaction to the unpredictable component of fiscal variables. A further approach is based on theoretically micro-founded DSGE models incorporating, in to varying degrees, New-Keynesian features, where equation parameters are either calibrated or estimated (or both). Generally, the state-dependence of fiscal multipliers and the non-linear effects of fiscal policy can be controlled for in a vector auto-regression framework by separating observations in different regimes based on a threshold variable (TVAR), while in a DSGE framework non-linearity can be an endogenous outcome produced by the theoretical formulation of the model. It is more complex to control for state-dependency in a structural macro-econometric model, which is linear by construction (fiscal policy has the same effect irrespective of the phase of the cycle), although, as shown in what follows, it is nonetheless possible with appropriate methodologies to retrieve useful information from the changing value of structural parameters and, in these models, to produce estimates of period-specific fiscal multipliers.

3 The Istat MeMo-It model

The aim of this section is to describe the main features of the Istat macro-econometric model MeMo-It (sub-section 3.1) and provide a preliminary overview of the characteristics of the fiscal multipliers of the model (sub-section 3.2). MeMo-It can be considered a small-scale traditional structural econometric model (SEM).[3] It is based on annual information and is mainly used for medium-term forecasts of the Italian economy as well as ex-ante evaluations of the impact of fiscal policy measures. The modelling approach adopted in MeMo-It is a mixture of both the London School of Economics (LSE) methodology on integrated and cointegrated systems (Hendry et al. 1984) and the Cowles Commission approach (referring to the specification and testing of structural macro-econometric models; Fair 1984, 2004, 2015). In order to merge theory and data, MeMo-It uses cointegration methods on dynamic subsystems to estimate theory-interpretable and identified steady-state relationships, specified in the form of error-correction models (ECM). The main features of the model can be summarized as follows.

First, MeMo-It makes explicit reference to empirical information in order to assess the data-admissibility of the theoretical constructs, while it does not assume explicit micro-foundations for the behavioral equations. Second, cointegration analysis is performed within the blocks of the model to check whether the theoretical model is a valid approximation of a steady-state equilibrium. Several equations are specified in the error-correction form, thus using the long-run cointegrating information in the data, but allowing for a more flexible short-run dynamics. Third, as regards the theoretical approach, MeMo-It is based on the New-Keynesian framework. The key assumption is that, in the short run, economic activity is mainly driven by the demand side, while in the long run the economic system converges to the potential output provided by the supply side of the model. Prices react to the output gap in order to account for the disequilibrium of supply and demand. Price changes cause a shift in demand-side variables and wages, which in turn affects income distribution and household consumption. Furthermore, higher inflation leads to lower competitiveness, lower exports, lower investment and lower output, thus crowding out the effect on GDP of non-government components in the long run. The next sub-section presents additional details on some specific issues relevant for the multiplier analysis.

3.1 MeMo-It settings

The macro-econometric model MeMo-It can be represented as a linear stochastic econometric model including *g* endogenous variables, *k* exogenous (or predetermined)

[3] The model used to perform the estimates reported in this paper consists of 60 stochastic equations and 80 identities. A comprehensive discussion of MeMo-It main characteristics is in Bacchini et al. (2013, 2015).

variables and a set of stochastic structural disturbance terms. In vector-matrix notation, the *structural form* can be written as follows,

$$\boldsymbol{A}y_t = \boldsymbol{A}_{\boldsymbol{1}}^* y_{t-1} + \cdots + \boldsymbol{A}_{\boldsymbol{p}}^* y_{t-p} + \boldsymbol{B}_{\boldsymbol{0}}^* x_t + \cdots + \boldsymbol{B}_{\boldsymbol{q}}^* x_{t-q} + \varepsilon_t \tag{1}$$

where $y_t = (y_{1t}, \ldots, y_{gt})'$ is the *g*-dimensional vector of endogenous variables, $x_t = (x_{1t}, \ldots, x_{kt})'$ is the *k*-dimensional vector of exogenous variables (which can include both stochastic and non-stochastic components); $\boldsymbol{A}$ is a (*g x g*) matrix assumed to be non-singular and represents the instantaneous relation between the endogenous variables; $\boldsymbol{A}_i^*$ and $\boldsymbol{B}_j^*$ are (*g* x *g*) and (*g* x *k*) coefficient matrices (*i*=1,...,*p*; *j*=0,...,*q*), and ε_t is a *g*-dimensional error vector. Each equation of the structural form can reflect a behavioral relation, a technological relation, or some other specific relation suggested by theory for the system under study.

The assumptions that define the statistical model are the standard ones. First, the structural disturbances are assumed to be randomly drawn from a stationary multivariate distribution with $E[\varepsilon_t|x_t] = 0$, so that ε_t and x_t are independent processes; furthermore, it is assumed that $E\left[\varepsilon_{ti}\varepsilon_{tj}'\middle|x_{ti}, x_{tj}\right] = \sigma_{ij}\boldsymbol{I}_{\boldsymbol{T}} = \boldsymbol{\Sigma}$ (*t*=1,2,...,*T*), where $\boldsymbol{\Sigma}$ is the positive-definite symmetric covariance matrix, and σ_{ij} the contemporaneous covariance between the disturbances of different equations. In addition, the structural errors are assumed to be uncorrelated over the sample, so that (1) determines the joint distribution of the variables y_t conditional on the predetermined variables x_t.

Given the nonsingularity assumption on $\boldsymbol{A}$**,** the structural form can be converted into *reduced form* by premultiplying with $\boldsymbol{A}^{-\boldsymbol{1}}$,

$$y_t = \boldsymbol{A}_1 y_{t-1} + \cdots + \boldsymbol{A}_p y_{t-p} + \boldsymbol{B}_0 x_t + \cdots + \boldsymbol{B}_q x_{t-q} + v_t \tag{2}$$

where $\boldsymbol{A}_i = \boldsymbol{A}^{-1}\boldsymbol{A}_i^*$, (*i*=1,...,*p*) $\boldsymbol{B}_j = \boldsymbol{A}^{-1}\boldsymbol{B}_j^*$, (*j*=0,...,*q*) are the matrices of reduced-form parameters, and $v_t = \boldsymbol{A}^{-1}\varepsilon_t$ is the reduced-form stochastic disturbance vector. In lag operator notation, the system in (2) can be represented as

$$\boldsymbol{A}(L)y_t = \boldsymbol{B}(L)x_t + \varepsilon_t \tag{3}$$

where both $\boldsymbol{A}(L) = (I_g - A_1L - \cdots - A_pL^p)$ and $\boldsymbol{B}(L) = (B_0 + B_1L + \cdots + B_qL^q)$ are polynomials in the lag operator of order *p* and *q*, respectively. These reduced-form equations are in general nonlinear in the structural parameters. Given the exogenous variables, they uniquely determine the probability distributions of the endogenous variables, the coefficients and the probability distributions of the stochastic disturbance terms.

Furthermore, some of the variables in y_t and x_t may be non-stationary. Since the essential relations between integrated variables are the cointegration relations, the dynamic simultaneous equations model in (1) can be written in vector error-correction form,

$$\boldsymbol{A}^*(L)\Delta y_t + \boldsymbol{B}^*(L)\Delta x_t + A(1)y_{t-1} + B(1)x_{t-1} = \varepsilon_t \tag{4}$$

where $A(1)y_{t-1} + B(1)x_{t-1}$ and $\boldsymbol{A}^*(L)\Delta y_t + \boldsymbol{B}^*(L)\Delta x_t$ denote, respectively, the long-run and short-run relations between y_t and x_t; the matrix $-A(1)$ denotes the measure of response of *g* jointly dependent variables to the deviation from the equilibrium in the previous period, with the latter defined by $y^*_{t-1} = -A(1)^{-1}B(1)x_{t-1}$.

As for the identification issue, Hsiao (1997) demonstrates that the identification of long-run equilibrium relations is not independent of the identification of short-run dynamics. There is only one set of conditions that simultaneously identify both the long-run equilibrium relations and short-run dynamics. And, notably, the necessary and sufficient condition for identifying the *g*-th equation of the error-correction representation is the usual rank condition as in traditional structural equation models. Therefore, identification can be studied equation-by-equation without loss of generality. Provided that certain conditions are satisfied,[4] the necessary and sufficient condition for identification of parameters of the *g*-th equation in system (1) is that the rank$[\boldsymbol{\Gamma}\varphi]$=*g*-1 (rank condition), where $\boldsymbol{\Gamma} = [\boldsymbol{A}, \dots, \boldsymbol{A}^*_p, \boldsymbol{B}^*_0, \dots, \boldsymbol{B}^*_q]$ is the matrix of all structural parameters, with the normalization restrictions imposed, and φ is the matrix of *r* further zero restrictions on the structural parameters (imposed prior to the estimation of the model), while no cross-equation restrictions and no restrictions on the covariance matrix are considered.

3.1.1 Estimation procedure

Once the *g*-th equation in system (1) satisfies the identification conditions, and prior to the estimation of the system of equations, it is preliminarily inspected by estimating parameters using the two-stage least squares (2SLS) estimator. This allows us to tackle the issue of simultaneity bias that arises because the RHS endogenous variables are correlated with the error term. Cointegration among the regressors transforms the cointegrated I(1) regressors into stationary regressors, but the correlations with the error terms do not disappear. Thus, the OLS are consistent only in the estimation of long-run relationships. By contrast, if the *g*-th equation is identified, then the order condition is also met, and a set of instruments is available to correct for simultaneity bias. Therefore, each equation of the system is estimated by two-stage least squares (Hsiao, 1997), which is a consistent though inefficient estimator (since the method does not account for the correlation of the disturbances across equations).

[4] Usual assumptions are the following: *i*) the matrix A is invertible; *ii*) $(1/T^2)\sum X'X$ converges in distribution to a non-singular random matrix; *iii*) there exist at most *g* linearly independent cointegration relations for the system.

In this framework, once identification conditions are satisfied, and the model specification is complete, the three-stage least squares (3SLS) estimator is applied to perform consistent estimates of the parameters of the structural form, as it accounts for both inconsistency of the OLS (due to the presence of endogenous regressors) and cross-equation correlation of the disturbance terms.

For specific applications of the econometric model (forecasting and ex-ante evaluation of the macroeconomic impact of policy measures), researchers might focus more on the reduced form of the model and, in such cases, the aim is to obtain consistent and efficient estimates of the reduced-form parameters. Under the assumptions set for identification, the conditions of both the Gauss-Markov theorem and the Least Squares Consistency Theorem are satisfied for the reduced-form system (2), so that the reduced-form parameters are observable and can be consistently estimated using the OLS estimator. But the estimates are not efficient since the method does not take account of overidentifying restrictions.

Alternatively, a well-known method is to derive the estimator of reduced-form coefficients from the consistent estimates of structural parameters (derived reduced-form estimates). This procedure leads to the possibility of a more efficient estimation of the reduced-form parameters as it accounts for both the sample information and the non-sample information in the form of restrictions on parameter matrices. Specifically, the 3SLS estimates of structural parameters are more efficient than the 2SLS estimates when at least one equation is overidentified (in the case of exact identification for all equations, both methods are equally efficient). The 3SLS derived reduced-form estimator is asymptotically more efficient than both 2SLS and OLS derived reduced-form estimators because the structural approach utilizes prior knowledge and this results in a smaller limiting covariance matrix. In what follows, the estimates of the reduced-form parameters are obtained by a non-linear combination of the 3SLS estimates of the structural coefficients.

3.1.2 Final form and dynamic multipliers

The matrix of reduced-form coefficients is particularly important as it represents what in economics is related to the concept of *multipliers*. From the reduced form in (2), the system can be solved for the endogenous variables by multiplying by $\boldsymbol{A}(L)^{-1}$ (assuming invertibility of $\boldsymbol{A}(L)$), so as to obtain the representation $y_t = \boldsymbol{\Pi}(L)x_t + \boldsymbol{A}(L)^{-1}v_t$, which is known as the *final form* of the system, with $\boldsymbol{\Pi}(L) = \boldsymbol{A}(L)^{-1}\boldsymbol{B}(L) = \sum_i \boldsymbol{\Pi}_i L^i$ (*i*=1,2,...). The coefficient matrices $\boldsymbol{\Pi}_i = (\boldsymbol{\pi}_{j,k,i})$, obtained from the expansion of $\boldsymbol{A}(L)^{-1}\boldsymbol{B}(L)$, denote the effect that a unit change in the exogenous variables has on the endogenous variables. Everything else held constant, a unit change in the *k*-th exogenous variable in period *t* induces a marginal change of $\boldsymbol{\pi}_{j,k,i}$ units in the *j*-th endogenous variable in period t + *i* (*i*=0,...,*h*,...,*n*). Therefore, the elements of the $\boldsymbol{\Pi}_i$

matrices are called *dynamic multipliers*: $\boldsymbol{\Pi}_0$ denotes the *impact multipliers* matrix (*i*=0), and $\boldsymbol{\Pi}_h$ is the *interim* multipliers matrix of lag *h* (*i*=*h*). If the variation in the level of a given exogenous variable is sustained over *h* periods, the effect on the endogenous variables is obtained through the matrix of *cumulative* interim multipliers $\boldsymbol{D}_h = \sum_{i=0}^{h} \boldsymbol{\Pi}_i$.

The estimation of these quantities can be obtained by substituting estimates $\widehat{\boldsymbol{A}}_i$ and $\widehat{\boldsymbol{B}}_j$ of the coefficient matrices in $\boldsymbol{A}(L)$ and $\boldsymbol{B}(L)$ (Lütkepohl, 2005).[5] The dynamic responses can also be computed numerically, as the difference between two solutions of a model, i.e. by comparing a perturbed solution of the model, in which the relevant shock is imposed, to a baseline (or control) solution. This latter approach has been adopted in this paper.

3.2 Preliminary evidence of possible shifts in fiscal multipliers in the crisis period

Multiplier analysis is a widespread approach to evaluating the properties of econometric models (such as, the size and the persistence of the effects of discretionary policy on output, the dynamic path of adjustment in the medium term), linking a model's performance to the theoretical assumptions embodied in its specification. As shown in sub-section 3.1, dynamic multipliers can be represented as nonlinear combinations of the final form model's parameters, and provide a summary impulse-response measure that is informative about the mechanics of transmission of the effects stemming from shocks to selected instruments (exogenous variables) to specific target variables.

Multiplier analysis involving MeMo-It has been carried out by Bacchini et al. (2013, 2015) for a restricted set of exogenous variables (government spending, government transfers to households, personal income taxes, consumption taxes), providing support to the new-Keynesian features of the model, with the short-term output responses to fiscal innovations, mainly related to demand-side shocks, tending to vanish in the long run.

This sub-section focuses on the short-term effects of discretionary policy changes, with the aim of providing preliminary evidence on possible shifts in the size of fiscal multipliers for the Italian economy during the crisis period (2008-2014). As discussed in Section 2, state dependency is a crucial issue in the evaluation of fiscal multipliers. MeMo-It, as all other structural econometric models, is characterized, by construction, by fiscal multipliers that do not depend on the state of the cycle at the moment the fiscal shock is imparted. It is however possible to enucleate from the behavior of the model information on the emergence of possible changes (nonlinearities) in the effects of fiscal shocks in the most recent period.

[5] The asymptotic distributions and the corresponding efficiency of estimates of the multiplier matrices are dependent upon the asymptotic distributions and the efficiency of the estimates of the reduced-form parameters. Therefore, the most efficient estimates of the multipliers are obtained by utilizing full-information-derived reduced-form estimates, cf. Schmidt (1973).

In what follows, the analysis is focused on the impact (first-year) multipliers, which are computed for several sample sizes. Fiscal multipliers are first estimated over the complete timespan from 1970 to 2014, which is therefore inclusive of the effects related to the financial crisis. As we know, it originated in the United States at the end of 2007 (December 2007 according to the NBER business cycle chronology) and thereafter it spread heterogeneously on a global scale. The European economies and Italy were immediately hit, falling into an initial severe recession at the start of 2008 and, after a short-lived recovery in 2010, slumping into a second recession related to the sovereign debt crisis that began in mid-2011.[6] The year 2007 can hence be considered as a watershed between "normal" and "crisis" times. On this basis, the 1970-2014 period was split into two subsamples: one that includes the information related to a "normal" sequence of expansions and recessions (1970-2007), and one that covers the years of the crisis (2008-2014).

A second set of fiscal multipliers was then computed for the 1970-2007 subsample, which by definition are unaffected by the factors related to the financial crisis. The 2008-2014 subsample represents the reference sample for obtaining direct estimates of the fiscal multipliers specific to the crisis period. Yet, the sample size is too small, so that MeMo-It cannot be used to obtain reliable inferences for the structural parameters and, therefore, to derive the corresponding fiscal multipliers. The approach to disentangle estimations of fiscal multipliers for such short period and the relevant results are presented in Section 5. Here, preliminary evidence of the state-dependence of fiscal multipliers is obtained by comparing the two sets of estimates, i.e. the output responses to fiscal shocks computed over the entire timespan (1970-2014) *vis-à-vis* those estimated for the pre-crisis period (1970-2007).

The impact (first-year) multipliers are computed in the form of output responses to permanent changes for a wide set of fiscal instruments.[7] The fiscal impulses are normalized so that the size of the discretionary shock represents a permanent increase in public spending or a permanent decline in public revenues (in the form of a cut in average tax rates) equal to 1 percent of nominal GDP. The change in fiscal variables is computed with reference to the baseline GDP level in the final year of the estimation sample (i.e., pre-stimulus nominal GDP in 2007 and 2014, respectively), and under the assumption of unchanged nominal short-term interest rates. Each set of fiscal multipliers (expressed in percentage points) is computed as the difference between two solutions of the model: the perturbed solution, in which a specific shock is imposed, and

[6] The Italian economy and the euro zone peaked in February 2008, according to both the OECD and Conference Board reference turning points series. Some authors date the beginning of the recession to the second half of 2008, e.g., Caggiano et al. (2015), Cimadomo and D'Agostino (2015).

[7] A permanent fiscal expansion may have smaller first-year multipliers compared with a temporary discretionary change, because the latter does not result in crowding out of private spending. But in the first year of the fiscal shock, the magnitude of both temporary and permanent multipliers can be very close, especially in medium-sized economies with no independent monetary policy. Indeed, the impact multipliers calculated using MeMo-It for both temporary and permanent discretionary fiscal changes turn out to be similar. This is in line with the results for the Italian economy reported in Barrell et al. (2012) and Carreras et al. (2016).

the baseline solution. Estimates are reported in Table 1: the impact (first-year) multipliers estimated over the complete sample (1970-2014) are presented in panel *a*; the multipliers pertaining to the pre-crisis period (1970-2007) are reported in panel *b*.

The main findings can be summarized as follows. First of all, the output responses to fiscal shocks estimated over the whole sample (Table 1, panel *a*) are similar to those reported in Bacchini et al. (2013)[8] as well as to the fiscal multipliers for Italy estimated using the NIGEM econometric model (Barrell et al. 2012; Carreras et al. 2016). By contrast, they are smaller than the fiscal multipliers (for some main fiscal shocks) underlying the models of some leading research institutions in Italy (see PBO 2017) and lower than the impact multipliers recently published by the Italian Ministry of the Economy and Finance (Felici et al. 2017). As for government investment, it is relevant to point out that the estimates of fiscal multipliers reported in Table 1 also include the output responses to shocks to investment grants. This broader definition of government investment (government capital expenditure and investment grants) is adopted hereafter.

Secondly, the whole-sample estimates are consistent with the empirical literature referred to in Section 2. The size of fiscal multipliers depends on the source of the exogenous impulse on GDP, with spending multipliers that tend to be larger than revenue multipliers in the short run. This is largely because spending items (i.e., government investment and consumption) directly impact aggregate demand. A special case is represented by investment grants, which affect aggregate demand through effects on firms' productivity, so that an increase in such transfers plays a similar role as a decrease in taxes.

Table 1 – Impact (first-year) multipliers – Full-sample and pre-crisis estimates

Fiscal variables	Point estimates	
	Full-sample estimates (1970-2014) *(a)*	Pre-crisis estimates (1970-2007) *(b)*
Intermediate consumption	0.602	0.504
Social transfers to households	0.150	0.112
Government investments	0.671	0.464
Households' labor income tax	0.160	0.122
Corporate income tax	0.017	0.018
Social security contributions	0.199	0.158
Consumption tax	0.074	0.056
Regional tax on economic activities	0.023	0.002
Excise duty on energy products	0.027	0.010

[8] The computations are performed using a version of MeMo-It model that, despite some minor changes, shares the same properties and performance as the Istat model.

The impact of this measure is expected to be even higher if the stimulus is targeted at credit-constrained firms. Analogously, short-run multipliers also tend to grow when social transfers are targeted at liquidity-constrained households, impacting aggregate demand through effects on households income and labor supply (Gechert et al. 2016).

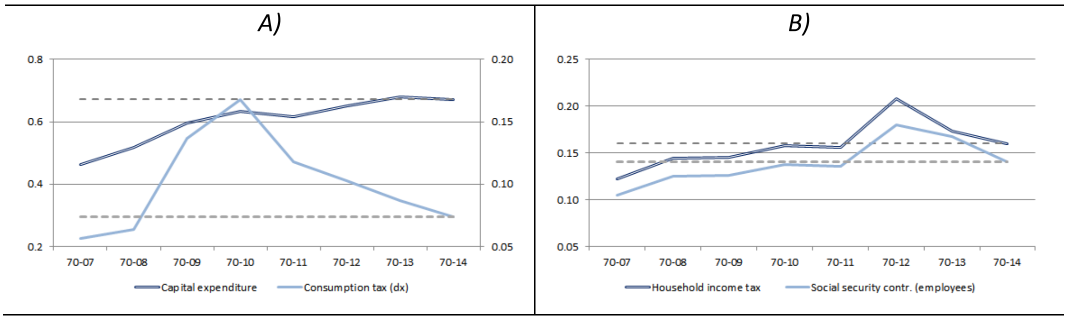
Furthermore, the comparison of fiscal multipliers reported in both panel *a* and panel *b* enables a preliminary assessment of the extent of state dependence, namely the heterogeneous output responses to fiscal shocks. It emerges that output responses to fiscal innovations estimated over the whole sample are generally larger in magnitude compared with the pre-crisis multipliers. The distance between the two sets of fiscal multipliers widens, especially for the expenditure categories (intermediate consumption and investment). On the revenue side, a difference in the size of multipliers between the two sample estimates emerges, especially for the exogenous variation in household income tax and consumption tax.

Such preliminary findings obtained through MeMo-It suggest that the crisis period in Italy (from 2008 onwards) was characterized by larger fiscal multipliers. When the difference in the size of fiscal multipliers in two subsamples is considerable, then such a gap should be mainly related to differences in the estimates of the corresponding structural parameters in the econometric model. This calls for further investigation into the stability of parameter estimates, since potential breaks in the time series could have occurred in the final part of the sample.

To obtain an initial insight into the issue of the stability of output responses to fiscal shocks, recursive estimates of the pre-crisis fiscal multipliers are computed over expanding data windows, i.e. 1970-2008, 1970-2009 up to 1970-2014 (full-sample). In performing this exercise, it should be borne in mind that the recursive estimates are feasible once the deterministic components (impulse and step dummies) needed to account for several factors that occurred from 2008 onwards (unexpected shifts in time series, country-specific institutional events) are removed from the baseline specification.[9] Therefore, possible instabilities detected through the recursive estimates could reflect the elimination of the deterministic component in selected equations of the model. Taking account of this, the inspection of recursive estimates provides some preliminary findings on the presence of potential parametric change. The results, reported in Figure 1, show that shifts from the 1970-2007 pre-crises estimates are detected in particular for the output responses to changes in capital expenditure (which tend to rise as the sample size increases until it approaches the full-sample estimates), changes in consumption tax (which initially rises as the sample size increases, overshooting the full-sample estimates and then subsiding) and, although to a lesser extent, changes in household income tax and social security contributions paid by employees.

[9] The deterministic components are removed in order to perform the recursive estimates in each data window. The final data window provides the full-sample estimates.

Figure 1 – Recursive estimates - Output responses to selected demand- and supply-side shocks (1)



(1) The dotted straight lines represent full-sample estimates (1970-2014).

Overall, the expenditure multipliers (to changes in intermediate consumption and in social transfers to households) do not signal a diverging path, as they fluctuate near the full-sample values. Disentangling whether the differences in fiscal multipliers in the two periods could reflect fluctuations around a long-term average or possible structural breaks requires thorough investigation. The issue of the stability of structural parameters in MeMo-It model is dealt with in Section 4.

4 Testing for structural breaks and re-estimation of fiscal multipliers

Estimation and inference in models with structural changes have received a great deal of attention in the theoretical econometrics literature, since economic policy decisions can be seriously biased by parameter instability in economic models. Controlling for structural breaks is a key issue, as the structural parameters of macroeconomic models could reflect the changing behavior that occurs in macroeconomic time series for a number of reasons, including policy changes and regime shifts (such as the global financial crisis and zero lower bound interest rates) as well as changes in institutional arrangements (the Maastricht Treaty may have induced a structural break in the way Italian government revenues and expenditures reacted to past levels of debt in 1993; cf. Caprioli and Momigliano 2011). Since fiscal multipliers are obtained as a non-linear combination of the estimates of structural-form parameters, the issue of instability is highly relevant.

Structural break analysis has a long tradition in the statistical and econometric literature. Perron (2006) provides an exhaustive survey, dating back to the seminal works by Quandt (1960), Gardner (1969) and Brown et al. (1975). More recently, a number of studies have developed different methodologies for endogenizing break dates. Andrews (1993) considers tests for parameter instability and structural change with unknown change points. Ghysels et al. (1998) provide predictive tests for structural change in models estimated applying the generalized method of moments and derive the limiting distribution of test statistics under both the null hypothesis and local alternatives. Bai and Perron (1998, 2003) consider theoretical and practical issues related to limiting distribution of estimators and test statistics in the linear model with multiple structural changes using the least squares principle. Hendry (1999) and Hendry et al. (2008) proposes an outlier-robust method known as "Dummy Saturation". It relies on pivotal statistics to ascertain the number of structural breaks and enables the joint treatment of outlying observations and structural breaks by saturating the initial model with different kinds of dummies.

The current specification of MeMo-It includes a parameterization of structural breaks in the form of discrete shift dummy variables, as they capture relationship modifications due to changes in the policy regime or exogenous shifts. From a modelling perspective, as the functional form adopted for the majority of single-equation econometric specifications is the Error Correction Model (ECM), the detection of structural changes should concern both the long-run cointegration relation and the dynamic short-term specification. A testing procedure is performed equation by equation and consists of the following steps. First, the discrete specification of structural changes, if present, is removed from each model equation. Second, the testing for structural changes is carried out to test hypothesis about parameter constancy in the long-run cointegrating relationship, investigating whether there exists a single cointegrating vector (absence of structural changes) or whether multiple long-run equilibria are identified, in which case each cointegration regime (before/after the estimation of a given change point) corresponds to a specific equilibrium. Third, the investigation for structural breaks focuses on the short-term dynamic specification of the regression model. It is performed considering the whole model specification inclusive of the cointegrating term, adjusted for the results of the testing for (multiple) structural changes if appropriate. A specific issue of this analysis is to assess whether taking into account for breaks in the long run enables a stable estimation of factor loading in each equation.

4.1 Methodology

In the application of the testing procedure, we distinguish two stages, each of them based on specific test statistics. Stage 1 investigates the presence of discrete shifts at multiple unknown break points in the cointegration relations. Stage 2 is devoted to the detection of structural breaks in the error-correction model.

4.1.1 Stage 1: testing for structural changes in the cointegrating relations

With regard to testing for structural breaks when variables are non-stationary, the case of interest is when variables are cointegrated. Accounting for parameter shifts is crucial in cointegration analysis since it normally involves long spans of data which are more likely to be affected by structural breaks.

In this paper, we apply the method proposed by Kejriwal and Perron (2008b, 2010), who provide a comprehensive treatment of the issues related to testing for multiple structural changes occurring at unknown dates in cointegrated regression models. The procedure builds on the framework of Bai and Perron (1998) but is extended to allow both I(0) and I(1) variables in the regression model. Kejriwal and Perron (2008b, 2010) consider a linear model with *m* structural changes and *h*=1,...,*m*+1 regimes

$$y_t = c_h + \boldsymbol{z}'_{ft}\delta_f + \boldsymbol{z}'_{bt}\delta_{bh} + \boldsymbol{x}'_{ft}\beta_f + \boldsymbol{x}'_{bt}\beta_{bh} + u_t \quad (5)$$

where the break points are treated as unknown. In this model, y_t is a dependent I(1) variable, $\boldsymbol{x}_{ft}$ $(p_f \times 1)$ and $\boldsymbol{x}_{bt}$ $(p_b \times 1)$ are vectors of I(0) variables, $\boldsymbol{z}_{ft}$ $(q_f \times 1)$ and $\boldsymbol{z}_{bt}$ $(q_b \times 1)$ are vectors of I(1) variables, where the subscript *b* represents "break" and the subscript *f* represents "fixed" (across regimes). To control for the presence of simultaneity bias, Kejriwal and Perron (2010) use the dynamic OLS estimator (DOLS), which consists in augmenting the OLS regression with leads and lags of the first-differences of the I(1) regressors, which can be selected on the basis of information criteria (Kejriwal and Perron, 2008a).

The final model for testing multiple structural breaks reduces to a pure structural change model

$$y_t = c_h + \boldsymbol{z}'_{bt}\delta_{bh} + \sum_{l=-p}^{p} \Delta\boldsymbol{z}'_{t-l}\Pi_l + v_t^* \quad (6)$$

assuming $(p_f = p_b = q_f = 0)$. Based on this framework, Kejriwal and Perron (2010) develop three testing procedures.

First, they propose a sup-*Wald* test of the null hypothesis of no structural break (*m* = 0) versus the alternative hypothesis of a fixed (arbitrary) number of breaks (*m* = *k*),

$$\sup F_T^*(k) = \sup_{\lambda \in \Lambda_\epsilon} \frac{SSR_0 - SSR_k}{\sigma^2}$$

where SSR_0 denotes the sum of squared residuals under the null hypothesis of no breaks, SSR_k denotes the sum of squared residuals under the alternative hypothesis of *k* breaks, σ^2 is a consistent estimator of long-run variance under the null of no structural change, $\lambda = (\lambda_1, \dots, \lambda_m)$ is the vector of break fractions defined by $\lambda_i = T_i/T$ for T_i the break date, (*i* = 1,..., *k*) and T the sample size.

Second, they consider a test of the null hypothesis of no structural break (*m* = 0) versus the alternative hypothesis that there is an unknown number of breaks, given some upper bound *M* (1 ≤ *m* ≤ *M*), so that the test is a double-maximum test ($UD\,max$) based on the maximum of the sup-*Wald* test for the null of no break versus *m* breaks (*m* = 1,...,*M*), defined as

$$UD\,max\, F_T^*(k) = \max_{1 \le m \le M} \sup_{\lambda \in \Lambda_\epsilon^m} F_T^*(\lambda, k)$$

Third, a sequential testing procedure (SEQ) is based on the estimates of the break dates obtained from a global minimization of sum of squared residuals, as described by Bai and Perron (1998). It consists in testing the null of k breaks against the alternative that one additional break exists, based on the test statistics

$$SEQ_T(k+1|k) = \max_{1 \le j \le k+1} \sup\nolimits_{\tau \in \Lambda_{j,\varepsilon}} T\{SSR_T(\hat{T}_1, \dots, \hat{T}_k) - SSR_T(\hat{T}_1, \dots, \hat{T}_{j-1}, \tau, \hat{T}_j, \dots, \hat{T}_k)\}/SSR_{T+1}.$$

Finally, Kejriwal (2008) also uses two additional procedures to select the number of breaks based, respectively, on the Bayesian Information Criterion (BIC) and on the modified Schwarz information criterion (Liu, Wu and Zidek 1997; LWZ hereafter). In this study, we mainly rely on the SEQ and LWZ procedures to assess the presence of structural breaks in cointegrating relations.

Kejriwal and Perron (2010) point out that their tests have power against a purely spurious regression, with the result that the null can be rejected even when no structural change is present and there is no cointegration. Kejriwal (2008) suggests using the above testing procedures in conjunction with cointegration tests allowing for structural changes in the parameters. We apply the Gregory-Hansen (1996) test (GH, hereafter) to confirm the presence of a cointegrating relationship when only one break is detected. As the GH test is designed to have power against the alternative of a single break in parameters, it may tend to accept the null of no-cointegration when the true data generating process exhibits cointegration with more than one break. For such cases, we use the Arai and Kurozumi (2007) residual-based test extended by Kejriwal (2008) by incorporating multiple breaks under the null hypothesis of cointegration against the alternative of no cointegration (AK, hereafter). The auxiliary equation for the AK test is augmented with leads and lags of first differences of I(1) regressors to account for potential endogeneity. In the case of a single break, the test statistic is given by

$$\tilde{V}_1(\hat{\lambda}) = (T^{-2} \sum_{t=1}^{T} S_t\,(\hat{\lambda})^2)/\hat{\Omega}_{11}$$

where $\hat{\Omega}_{11}$ is a consistent estimate of the long run variance of v_t^* in equation (6), $\hat{\lambda} = (\hat{\lambda}_1, \dots, \hat{\lambda}_m)$ is the vector of break fractions defined by $\hat{\lambda}_i = \hat{T}_i/T$, where the break dates $\hat{T}_i$ (i = 1,..., k) are obtained using the dynamic algorithm of Bai and Perron (2003).[10]

[10] Several cointegration tests can be used to complement this analysis. In the case of a single break in the cointegration relation, one can refer to the test for the null of cointegration proposed by Shin (1994), Arai and Kurozumi (2007) (both applied, for example, in Dülger 2016), Carrion-i-Sylvestre and Sansò (2006) (see Beyer et al. 2009). Hatemi-J (2008) provides a test for the case of two unknown regime shifts in the long-run relation, recently applied in Bagnai et al. (2017). Maki (2012) proposes a cointegration test allowing for an unknown number of breaks.

4.1.2 *Stage 2: testing for structural changes in the error-correction model with endogenous regressors*

Following the estimation of structural changes in the cointegrating relationships, the analysis of parameter instability concerns the error-correction model as a whole, which consists of both the cointegrating term, eventually corrected for structural breaks, and of the related short-term specification. The aim is to identify potential break points in the short-term specification of the model and in the factor loading once the structural instability in the long-term component has been controlled for.

The existence of endogenous regressors has to be explicitly considered in the estimation of break points in the single equation model as a whole. By ignoring endogeneity, breaks in the reduced form can potentially bias inferences about breaks in the structural equation. As mentioned in Section 3, each equation in the MeMo-It model is estimated, as a preliminary step, via 2SLS in order to account for potential endogenous regressors . Therefore, in order to control for structural breaks, we focus on tests for structural instability within the IV framework.

In what follows we apply the technique proposed by Perron and Yamamoto (2015; PY hereafter).[11] They show that, even in the presence of endogenous regressors, it is preferable to estimate the break dates and test for structural changes using the usual OLS-based framework. They point out some advantages in adopting this framework. First, using OLS delivers consistent estimates of the break fractions and, in the majority of cases, improves the efficiency of the estimates and the power of the tests. The reason is that the IV-based framework involves as regressors the projection of the original regressors on the space spanned by the instruments, which have less quadratic variation than the originals. Second, the OLS procedure avoids loss in efficiency, which can be pronounced when the instruments are weak. Therefore, the objective is to perform IV-based regression but conditioning on the estimates of the break dates obtained using the OLS-based procedure.

Under some assumptions, Perron and Yamamoto (2015) demonstrate that the probability limit of OLS estimates $\hat{\delta}$ can be denoted as

$$\delta^{*} = \delta^{0} + [(Q_{XX}^{1})^{-1}\varphi_{1}, \dots, (Q_{XX}^{m+1})^{-1}\varphi_{m+1}]$$

where δ^{*} is the limit value of OLS estimates, δ^{0} is the true value of the coefficient vector and $(Q_{XX})^{-1}\varphi$ is the bias term. They point out that in the vast majority of cases it is preferable to estimate the break dates in the structural form by applying the OLS-based procedure (Bai and Perron 1998, 2003), which outweighs the IV-based procedure (Hall

[11] As for the alternative approaches in the IV framework, Hall et al. (2012) and Boldea et al. (2012) extend the OLS approach of Bai and Perron (1998) and develop a hypothesis testing procedure for structural breaks that is (asymptotically) valid in the 2SLS framework. Hall et al. (2015) establish conditions under which the information criteria yield consistent estimates of the number of breaks when employed in the second stage of a 2SLS procedure, with breaks in the reduced form taken into account in the first stage.

et al., 2012) when a change in the structural form occurs. They point out some exceptions in which IV is marginally better than OLS. First, the acceptance of the null can occur when the change in δ^0 across h regimes (h=1,...,m+1) is exactly offset by the change in the bias term (knife-edge case, $\delta_h^* = \delta_{h-1}^*$ and $\delta_h^0 \neq \delta_{h-1}^0$). A second case is when the OLS-based methods reject the null hypothesis of no change in the structural form when none occurs ($\delta_h^* \neq \delta_{h-1}^*$ and $\delta_h^0 = \delta_{h-1}^0$), due to possible changes in the marginal distributions of the regressors or the correlation between the errors and regressors. Overall, in performing the tests, some care is required: the source of the rejection should be carefully assessed by evaluating the values of the changes across segments in both the structural parameters and the bias terms, and verifying that the change in the probability limit of the parameter estimates is not due to a change in the bias term. The latter can be computed after the IV estimation of the structural model based on the OLS-based estimates of the break dates.

4.2 Empirical results

This sub-section presents the empirical results of the testing procedure for multiple structural breaks. First, we discuss the results of the test for the presence of multiple structural changes in the long-run relationship specific to each MeMo-It equation. They are obtained by applying the method proposed by Kejriwal and Perron (2010) with the following settings: both the intercept and the slope are permitted to change, heteroskedasticity across sub-samples is allowed for, the trimming value is set to 15% so that the maximum number of structural changes allowed under the alternative hypothesis is 5. The leads and lags in DOLS specification are selected based on the Schwarz information criterion as DOLS results are sensitive to the exact number of leads and lags in small samples (Kejriwal and Perron, 2008a). Long-run variances are estimated with a quadratic spectral kernel and automatic bandwidth selection (Andrews, 1991). As a strategy to identify the number of changes and the corresponding break dates, we first check if any structural break is present (based on the double maximum break test, *UDmax*) and, if evidence in favor of a break is found, we consider the findings obtained by applying the sequential procedure SEQ and the information criteria BIC and LWZ.

Overall, the testing procedure finds evidence of at least one structural instability in many of the cointegrating relationships. In the majority of cases, only a single structural change is selected. Results are reported in Table 2, that presents the results of the stability tests as well as the number of breaks selected by both the sequential procedure and the information criteria (results refer to the number of breakpoints for which the null of no structural change is rejected at a 1% significance level).

Table 2 – Structural break tests (1)

Cointegrating relationships (2)	UDmax		SEQ	BIC	LWZ	T_1	T_2	T_3
Equation 1	363.5	***	3	4	3	1962	1988	2001
Equation 2	54.6	***	2	3	1	1981		
Equation 3	11.3	***	0	1	0			
Equation 4	46.8	***	1	1	1	2004		
Equation 5	113.6	***	2	1	1	1986	1993	
Equation 6	115.8	***	3	3	2	1979	1987	
Equation 7	42.9	***	1	1	1	1994		
Equation 8	4.3		0	0	0			
Equation 9	50.8	***	1	1	1	1987		
Equation 10	113.6	***	1	1	1	1989		
Equation 11	8.8		0	1	0	1993		
Equation 12	90.2	***	0	3	1			
Equation 13	9.2		0	0	0			
Equation 14	110.5	***	1	1	1	1993		
Equation 15	102.7	***	1	2	1	1993		
Equation 16	8.1		0	1	0			
Equation 17	41.8	***	3	2	2	1984	1995	
Equation 18	3.9		0	0	0			

(1) Specification of the SEQ procedure: the number of regressors allowed to change are 2 (intercept and slope, regime shift model) except for equations 4, 5, 7, 14, 15, which include a number of variables greater than 2; the number of leads and lags for the DOLS estimation is specific to each equation and set according to the Schwartz information criterion; the trimming value is set equal to 0.15 or higher depending on the length of the period; the maximum number of structural changes allowed is 5 for equations 1, 2, 6, 12 while for the other relations the number of breakpoints is set equal to the maximum number of breaks possible; UDmax is the tests of L+1 vs. L sequentially determined breaks; SEQ denotes the number of breakpoints selected according to the Kejriwal and Perron (2010) sequential procedure; BIC and LWZ denote the number of structural changes detected according to the information criteria. T_1, T_2 and T_3 denote the break dates selected by both the SEQ procedure and the LWZ criterion; for equation 3 the break date is selected in accordance with the LWZ criterion; for equation 5 the break dates are detected through the SEQ procedure; for equation 11 the break date is defined in accordance with the BIC criterion. Critical values are from Kejriwal and Perron (2010). ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. – (2) Cointegrated relations. Equation 1: private consumption and disposable income; equation 2: US and Italian stock market indices; equation 3: housing investment and disposable income; equation 4: capital stock (non-residential construction), output and user cost of capital ; equation 5: capital stock (equipment), output and user cost of capital; equation 6: ICT investment and gross operating surplus; equation 7: exports, world demand and real effective exchange rate; equation 8: imports (services) and domestic demand; equation 9: imports (non-fuel), domestic demand and relative prices; equation 10: imports (fuel), domestic demand and relative prices; equation 11: import prices (non-fuel) and world manufacturing prices; equation 12: implicit price deflator of investment in housing and overall investment implicit price deflator; equation 13: per capita wages in the public and in the private sector of the economy; equation 14: FTEs, output and per capita compensation of employees (private sector); equation 15: total employees (LFS) and FTEs in both the private and public sector; equation 16: oil demand and domestic demand; equation 17: compensations of employees in the household institutional sector and in the economy as a whole; equation 18: household current transfers (received) and government social expenditure.

The cointegrating relation between private consumption and disposable income (equation 1) is affected by three structural breaks (in 1962, 1988 and 2001) according to both the sequential procedure and the LWZ information criterion. In the remaining cases of multiple structural breaks, the stability tests detect a distinct number of breaks and/or a different timing of the break dates for the same relationship. In the case of the relation between ICT investment and real operating surplus (equation 6), the SEQ procedure selects three breaks while the LWZ criterion detect only two structural changes. The same finding applies to the relation between compensations of employees in the household institutional sector and the economy as a whole (equation 17).

Regarding the relationship between the US and the Italian stock market indices (equation 2), the sequential procedure identifies two breaks while only one change is identified by the modified Schwarz criterion. Similar results are found in testing the long-run relation between capital stock (equipment), real output and user cost of capital (equation 5).

For some cointegrating relationships, both the sequential procedure and the information criteria select only one break point occurring at the same date. This evidence suggests the existence of two-regime long-run models, with the break dates that are selected in a period ranging from the early 1990s to the mid-2000s. We do not investigate at this stage on the sources of shocks, which identification would require a specific analysis, that is beyond the scope of the present paper. As a general comment it can be pointed out that they seem to be related to significant economic and institutional occurrences. As an example, in the model for the long-term development of labor demand in the private sector, output and per capita labor income (equation 14), a break date is identified in 1993, which is the year of the so-called “July agreement”, which established a closer link between anticipated inflation, productivity and wage negotiations. In the long-run relation between exports, world demand and real effective exchange rates (equation 7), the stability tests detect a break in 1994, a date close to the 1992 exchange rate crisis and the sharp devaluation of Italian lira.

Finally, the tests do not suggest any instability in the cointegrating relationships linking investment in dwellings, disposable income and interest rates (equation 3), the implicit price deflator of investment in housing and the overall investment price deflator (equation 12), per capita wages in both the public and the private sector of the economy (equation 13), oil demand in Italy and domestic demand (equation 16), household current transfers (received) and government social expenditure (equation 18).

A further step in the inference is to confirm that the rejection of the null of stability is obtained by the presence of cointegration with multiple breaks, since the above stability tests reject the null of cointegration even when the regression is spurious, i.e. when no structural change is present and there is no cointegration. To ensure the existence of cointegration, a number of additional cointegration tests are used as confirmatory tests.[12] First, the residual-based test of Phillips and Ouliaris (1990, PO hereafter) for the null of no cointegration is used (the alternative hypothesis is cointegration without breaks; the lag length to account for serial correlation is selected using the Schwarz criterion). Second, concerning the one-break cointegration tests, the Gregory and Hansen (1996) test of the null of no cointegration against the alternative of cointegration with a single break is adopted. The test statistics (Z^*_α, Z^*_t and ADF^*) are estimated allowing for changes in both the constant and the slope (regime shift model). A similar setting is also used for the AK test for the null of cointegration with one structural break against the alternative of no cointegration. Third, the AK cointegration test with multiple structural breaks, as extended by Kejriwal (2008) to incorporate multiple breaks under the null hypothesis of cointegration, is applied according to the number of breaks estimated by both the SEQ and the information criteria. As an example, for equation 2 in Table 2, the sequential procedure selects two breaks while the LWZ criterion estimates a single break. Therefore, the one-break cointegration tests (GH and AK) and the two-breaks test (the extended AK test) are used to verify cointegration. Results are reported in Table 3.

Table 3 – PO, GH and AK cointegration break tests (1)

Cointegrating relationships[2]	PO		GH						AK one-break		AK two-breaks				AK three-breaks			
	Z_t		ADF		Z^*_α		Z^*_t		$V(\lambda_1)$	λ_1	$V(\lambda_2)$		λ_1	λ_2	$V(\lambda_3)$	λ_1	λ_2	λ_3
Equation 1	-8.38		-3.69		-20.15		-3.78		0.10	0.14					0.03	0.17	0.59	0.79
Equation 2	-7.33		-5.25	***	-29.41		-4.37		0.09	0.41	0.06		0.43	0.88				
Equation 3	-24.02	***	-4.39		-19.98		-3.59		0.07	0.41								
Equation 4	-2.37		-4.46		-23.06		-4.04		0.05	0.83								
Equation 5	-8.36		-4.42		-26.62		-4.78		0.04	0.69	0.06		0.17	0.41				
Equation 6	-10.28		-4.31		-21.08		-4.03		0.07	0.21	0.06		0.19	0.38	0.04	0.24	0.38	0.64
Equation 7	-3.78		-5.1		-23.79		-4.05		0.06	0.31								
Equation 8	-39.83	***	-11.36	***	-53.04	**	-11.6	***	0.09	0.15								
Equation 9	-10.41		-5.59	***	-28.97		-5.02	**	0.06	0.19								
Equation 10	-20.03	**	-5.02	**	-30.23		-5.55	***	0.08	0.19								
Equation 11	-10.26		-5.76	***	-25.72		-4.76	*	0.12	0.81								
Equation 12	-17.65	**	-4.41		-25.77		-4.15		0.10	0.81					0.02	0.48	0.74	0.88
Equation 13	-14.53	*	-3.78		-22.62		-4.10		0.12	0.14								
Equation 14	-3.96		-5.12		-26.44		-4.52		0.05	0.31								
Equation 15	-8.72		-6.61	***	-39.21		-6.61	***	0.09	0.23	0.12	**	0.20	0.75				
Equation 16	3.38		-3.95		-23.46		-4.12	*	0.04	0.26								
Equation 17	-8.03		-4.31		-23.63		-4.03		0.04	0.44	0.04		0.36	0.64	0.03	0.18	0.49	0.64
Equation 18	-19.54	**	-6.55	***	-31.03		-6.83	***	0.07	0.12								

(1) PO reports the Phillips and Ouliaris (1990) cointegration test Z-statistic. The null hypothesis is that series are not cointegrated. GH reports the Gregory and Hansen (1996) test statistics for the null hypothesis of no cointegration against the alternative of cointegration with one breakpoint. Critical values are taken from Gregory and Hansen (1996). AK one-break reports the Arai and Kurozumi (2007) test statistics of the null of cointegration with one break. Critical values are from Arai and Kourozumi (2007). AK two-breaks and AK three-breaks refer to the test statistics of the AK cointegration test extended to the case of multiple breaks according to Kejriwal (2008). Critical values are obtained by simulations using 500 steps and 2000 replications. ***, ** and * denote significance at 1%, 5% and 10% levels, respectively. – (2) For the description of equations, see note to Table 2.

[12] An alternative strategy to confirm the presence of cointegration is to test for cointegration in each of the *m*+1 regimes, given *m* breaks selected through the stability tests, by applying the Johansen (1995) cointegration test; cf. Haug et al. (2006).

According to the empirical results, the PO test rejects the null of no cointegration for all equations for which the above stability tests do not select any breakpoint (at the 1% level for equation 3 and 8, at the 5% level for equations 10, 18, at the 10% level for equation 13). The null of no cointegration is also rejected (at the 5% level) for equation 12, consistent with the SEQ procedure although the LWZ stability criterion detected the presence of one structural change. For the other relationships, the existence of cointegration can only be supported in the presence of at least one structural break.

The GH and AK tests confirm the existence of one breakpoint in all cointegration relationships for which only one break is selected by both the SEQ procedure and the information criteria. This is the case for equations 9 and 10, for which the GH and AK tests suggest cointegration with one break at the 1% level. Cointegration with one break is only supported by the AK test for equations 4, 7 and 14 (the null cannot be rejected at the 1% level) while, for the same relationships, the GH test cannot reject the null of no cointegration.

In two cases, testing results are inconclusive. For equation 11, the double maximum test (UDmax) signals the absence of any break date, but the assumption of cointegration without breaks is not supported by the PO test; furthermore, the GH and AK tests are consistent with the results of BIC criterion, supporting the evidence of cointegration with one breakpoint. For equation 16, we conclude in favor on no structural breaks because the finding of cointegration with one break is weaker (at 10% level for the GH test, at the 1 percent level for the AK test).

The results of cointegration tests allowing for multiple changes are also considered. They provide additional evidence in selecting the number of breakpoints when the results of stability tests are mixed. The results of the GH and AK tests confirm the presence of cointegration with a single breakpoint for equation 2 (consistent with the LWZ criterion), although the null of cointegration with two-breaks cannot be rejected according to the extended AK test. The results for equation 15 are also in favor of one breakpoint, since the extended AK test with two breaks rejects the null at the 5% level.

In the remaining relationships, the presence of cointegration with multiple structural breaks is generally confirmed since the extended AK test cannot reject the null hypothesis, while the results of GH and one-break AK tests contradict each other. This applies to equation 5, as the extended AK test cannot reject the null of cointegration with two breaks (confirming the results of SEQ procedure), and to equations 6 and 17 (consistent with the LWZ criterion); for equation 1, the AK test accepts the null of cointegration with three structural changes (confirming the findings of SEQ and LWZ tests). Overall, the empirical analysis provides evidence of cointegration when multiple structural breaks are properly detected and modelled.

Based on the above findings, the cointegrating relations of MeMo-It are modified accordingly. We generally refer to the structural changes estimated by both the SEQ procedure and the criterion LWZ. We take into account the breakpoints selected through the LWZ criterion for the equations 2, 6, 12 and 17, the break date detected by the SEQ procedure for equation 5, and the structural change selected using the BIC criterion for equation 11.

As a subsequent step, the testing procedure sought to investigate the stability of the parameters in each model equation. For the subset of equations specified as an error-correction model, the investigation concerns both the parameters of the differenced explicative variables and the adjustment coefficient (loading factor). The assessment is performed by applying the Perron-Yamamoto (2015) method, which is intended to detect structural changes in linear regression models with endogenous regressors. It consists of two stages.

The first stage is devoted to inferring potential structural instability in both the reduced and structural forms of each model equation. This analysis is performed in an OLS framework so that the Bai and Perron (1998, 2003) testing procedure applies to deliver estimates of the break dates. As for the reduced forms, the investigation concerns the explicative variables, which are treated as endogenous, and, specifically, the stability of the relationships linking each endogenous RHS variable to the exogenous covariates and the excluded instruments. If any break point is found, the reduced form equation is specified accordingly and re-estimated to obtain the corresponding break-adjusted predicted values. These are needed in the next stage of the testing procedure. This strategy also applies to the long-run relationships, which are considered as endogenous. In this case, the reduced forms refer to the break-adjusted cointegrated relationships, whose structural changes are estimated through the Kejriwal and Perron procedure (instead of the Bai and Perron method). The OLS-based procedure also applies to the structural forms, which entail as regressors all the original covariates, including the endogenous ones (by contrast, in the usual IV framework, regressors are the projection of the originals on the space spanned by the instruments).

In the second stage, the objective is to evaluate the relevance of the break dates detected in the structural equations by assessing if the changes across regimes in the estimated structural parameters reflect genuine instability or if they are mainly due to a change in the bias term (or if a change in the *true* parameters is offset by any change in the bias factor). This analysis is performed within the IV framework but conditioning on the estimates of the break dates detected through the OLS-based procedure.

The evidence of the Perron-Yamamoto method points out that the large majority of the equations in the MeMo-It model do not show any significant evidence of structural changes. The main results can be summarized as follows.

First, the PY procedure finds instability in the structural forms in 26 out of 60 equations (the OLS-based sup-Wald test rejects the null of stability). Within this subset of models, the number of estimated reduced forms is equal to 110 (46 out of 156 covariates are exogenous) and, of these, more than half are found to be unstable. Structural breaks in both the structural and reduced forms involve 17 out of the 26 equation specifications (and about 24 percent of the unstable reduced forms), in the large majority of cases with distinct break dates in the structural and the reduced forms. In the remainder of the IV specifications with instability in the structural equation, the RHS variables as a whole are strictly exogenous (the covariates are instrumented by themselves). In this case, the second stage of the PY procedure does not apply and the stability of structural form is obtained by re-specifying the models in accordance with the results of the Bai and Perron methodology (1998, 2003). The same applies to the reduced forms with structural changes entering models whose structural equation is stable (about 34 percent of unstable reduced forms).

Second, the adjustment for break dates also applies to the unstable reduced forms in models with instability in the structural form. Such a correction allows us to assess if the breaks estimated in the structural equation correspond to a *true* instability or not. Once the instability of the reduced form is accounted for, the 2SLS estimation consists in running an IV regression but conditioning on the estimates of the break dates (in the structural form) obtained using the Bai and Perron (1998) procedure. The ultimate objective is to evaluate the values in both the structural parameters and the bias terms and verify that the change of the OLS parameter estimates across regimes is not due to a change in the bias factor. The main results are reported in Table 4.[13] According to the empirical findings, the structural instabilities detected in 9 out of 17 structural forms do not genuinely reflect a change in the structural parameters, because the main source of variation in parameter estimates across regimes is represented by the change in the bias term. Therefore, the adjustment for break dates of the structural form applies to 8 out of 60 model equations.

Finally, the majority of the cointegrated relationships detected as unstable through the Kejriwal and Perron (2010) methodology (9 out of 13 cases) enter equations with a stable structural form. Once the instability in the long-run equilibrium relationships is accounted for, the coefficient of the short-run dynamic adjustment mechanism is found to be stable in all ECM specifications. It is worth pointing out that, as for the instability of the cointegrating relationships, the large majority of break dates are detected before 2008. These findings suggest that the change in the size of fiscal multipliers in the crisis period is not related to spurious factors, e.g. to the presence of break dates close to the inception of the financial crisis.

[13] Full results of the Perron-Yamamoto testing procedure are not reported for space considerations. They are available upon request.

Table 4 – IV subsample estimates of the unstable structural forms (1)

Equations[2]	Break dates	Reduced forms	Change in structural parameters				Change in bias term			
			m_2-m_1	m_3-m_2	m_4-m_3	m_5-m_4	m_2-m_1	m_3-m_2	m_4-m_3	m_5-m_4
Eq. 1	1	Rf1	0.195				-0.057			
Eq. 2	2	Rf1	10.266	-2.023			-30.444	18.357		
		Rf2	9.466	-1.038			-9.593	1.065		
Eq. 3	1	Rf1	-0.957				1.268			
Eq. 4	1	Rf1	0.085				0			
Eq. 5	3	Rf1	0.528	-1.176	0.901		-0.117	0.880	-0.937	
Eq. 6	1	Rf1	0.522				-0.107			
Eq. 7	1	Rf1	-0.352				1.052			
		Rf2	0.004				0.015			
Eq. 8	1	Rf1	-0.073				0			
		Rf2	-3.168				0			
Eq. 9	1	Rf1	-0.608				0.002			
Eq. 10	1	Rf1	-0.966				-0.272			
Eq. 11	1	Rf1	-0.407				-0.026			
Eq. 12	1	Rf1	0.478				-0.055			
		Rf2	1.351				-0.264			
Eq. 13	1	Rf1	-1.250				0.785			
		Rf2	0.089				-0.147			
Eq. 14	2	Rf1	0.003	-0.001			0	0		
		Rf2	-0.344	-0.682			0.297	0.523		
Eq. 15	1	Rf1	-3.586				0.826			
		Rf2	-3.630				0.914			
		Rf3	3.746				-0.962			
Eq. 16	4	Rf1	0.001	-0.016	0.017	0.001	0	0	0	0
Eq. 17	2	Rf1	0.138	-0.126			-0.112	0.191		

(1) Break dates report the number of *m* breakpoints (and *m*+1 regimes) estimated for each structural equation. Reduced forms are the number of estimated reduced forms for each model equation. Change in structural parameters reports the change in the coefficient estimated for each reduced form across the *m*+1 regimes. Change in bias is the estimated change in bias from each equation with the selected number of breakpoints in the structural equation. – (2) Equation 1: residential investments; equation 2: ICT and R&D investments; equation 3: exports of goods and services; equation 4: value added implicit deflator; equation 5: implicit price deflator for imports of services; equation 6: implicit price deflator for imports of goods and services; equation 7: implicit price deflator for gross fixed capital formation; equation 8: implicit price deflator for residential investment; equation 9: harmonized index of consumer prices, energy; equation 10: harmonized index of consumer prices; equation 11: wage per employee in the public sector; equation 12: labor force participation rate, females; equation 13: total labor force; equation 14: Government average cost of debt, % rate; equation 15: gross operating surplus, household institutional sector; equation 16: net property income, household institutional sector; equation 17: net property income, rest of the world.

4.3 Correction for structural breaks and significance tests of differences in multipliers

The MeMo-It model specifications are revised on the basis of the results of the stability tests. Furthermore, the model is estimated by using stochastic simulation, so that estimates of the fiscal multipliers are obtained along with their standard errors for assessing statistical significance.

In performing the stochastic simulation, the model is solved repeatedly for different draws of the stochastic components. As stated in Section 3, we assume that the vector of error terms, corresponding to the *m* stochastic equations of the model, is distributed as a multivariate normal $N = (0, \Sigma)$, where the covariance matrix Σ can be estimated as $\widehat{\Sigma} = (1/T)\widehat{U}\widehat{U}'$ and $\widehat{U}$ is the matrix of estimated residuals within a specified sample. For each repetition, the innovations of the model equations are generated by drawing a set of random numbers from $N = \left(0, \widehat{\Sigma}\right)$ distribution, that are scaled to match the variance-covariance structure of the system. We do not account for parameter uncertainty and the model coefficients are held fixed to the values of the (deterministic) dynamic simulation. Therefore, for each simulation period t, J=1,000 repetitions are performed, each characterized with a specific draw of the error terms from the $N = \left(0, \widehat{\Sigma}\right)$ distribution and fixed coefficient estimates (Fair, 2003, 2013).[14]

As the final aim of the stochastic simulation is to compute estimates of the sample variances of fiscal multipliers, for each repetition two scenarios are considered, one corresponding to the baseline model solution (no-stimulus scenario) and an alternate scenario, consisting in the specification of a permanent change in the values of a reference exogenous fiscal variable corresponding to 1 percent of nominal GDP in the pre-stimulus year of the simulation (such a measure is not balanced through changes in revenue or spending variables to preserve budget neutrality). The two scenarios are solved together stochastically to ensure that the same set of random shocks is used in both cases. For each repetition, one obtains a prediction of each endogenous variable. Let y_{it}^{j} denotes the value of the *j*th repetition of variable *i* for period t. For J repetitions, the stochastic simulation estimate of the expected value of variable *i* for period t is $\bar{\mu}_{it} = \frac{1}{J}\sum_{j=1}^{J} y_{it}^{j}$ and the estimate of the corresponding variance is $Var(\bar{\mu}_{it}) = \frac{1}{J}\sum_{j=1}^{J}\left(y_{it}^{j} - \bar{\mu}_{it}\right)^2$. The simulation estimate of the standard error is calculated by using the sample standard deviation formula, $\{Var(\bar{\mu}_{it})\}^{1/2}$.

As for output multipliers, for each repetition and for each period t, we compute the differences between the solution values for (log transformed) real output in both the baseline and the alternate scenario, denoted as $y_{t,B}^{j}$ and $y_{t,A}^{j}$, respectively. The fiscal multiplier for repetition *j* in period t is computed as $m_t^j = \log(y_{t,A}^{j}) - \log(y_{t,B}^{j})$. The mean of the quantities m_t^j over J repetitions provides the expected value of the stimulus to output in period t (so that the average is close to the fiscal multiplier obtained from the deterministic simulation). The corresponding stochastic simulation estimates of the variances and standard errors are also computed.

[14] Using bootstrapped innovations by drawing randomly (with replacement) from the set of actual residuals may be more appropriate than normal random numbers in cases where the equation innovations do not seem to follow a normal distribution. In this exercise, the results obtained do not qualitatively change using innovations generated by bootstrapping.

Within this methodological framework, we are able to replicate the exercise described in Section 3. The revised macro-econometric model is estimated for both the full-sample (1970-2014) and the period prior to the outbreak of the financial crisis (1970-2007). Based on the results specific to each sample estimate (estimated parameters and residuals), a 1-year out-of-sample stochastic simulation is performed to compute the impact multipliers and the related standard errors. Results are reported in Table 5 (multiplier estimates over the whole sample) and Table 6 (multiplier estimates for the pre-crisis period).

The comparison with the estimates discussed in Section 3 (Table 1) points out the following. First, the 1-year fiscal multipliers presented in Table 5 are generally larger than the multipliers estimated not accounting for the structural breaks. Such differences mainly concern the output responses to changes in government spending (intermediate consumption, social transfers to households, investment). A size increase is also observed, albeit to a lesser extent, for the output responses to a shock in households' labor income tax rate. Conversely, when the pre-crisis sample size is considered, the government expenditure multipliers reported in Table 6 are lower compared with their estimates not accounting for the break dates.

Table 5 – Fiscal multipliers from the break-corrected macroeconomic model. Full-sample estimates. (1)

Fiscal variables	Point estimates					
	1-year		2-year		3-year	
Intermediate consumption	0.864	***	0.946	***	0.987	***
	(0.056)		(0.072)		(0.054)	
Social transfers to households	0.180	***	0.497	***	0.554	***
	(0.013)		(0.030)		(0.027)	
Government investment	0.936	***	1.051	***	1.090	***
	(0.068)		(0.094)		(0.115)	
Production grants	0.033		0.052		0.109	***
	(0.163)		(0.063)		(0.045)	
Households' labor income tax	0.194	***	0.554	***	0.653	***
	(0.011)		(0.029)		(0.029)	
Corporate income tax	0.021	***	0.055	***	0.082	***
	(0.001)		(0.001)		(0.005)	
Social security contributions	0.099	***	0.323	***	0.424	***
	(0.023)		(0.035)		(0.042)	
Consumption tax	0.087	***	0.307	***	0.395	***
	(0.006)		(0.025)		(0.026)	
Regional tax on economic activities	0.034		0.115	***	0.142	***
	(0.094)		(0.032)		(0.021)	
Excise duty on energy products	0.035		0.163	***	0.270	***
	(0.058)		(0.024)		(0.029)	

(1) Standard errors in parenthesis. ***, **, * denote significance at 1%, 5% and 10% levels, respectively.

Table 6 – Fiscal multipliers from the break-corrected macroeconomic model – Pre-crisis estimates (1)

Fiscal variables	Point estimates					
	1-year		2-year		3-year	
Intermediate consumption	0.523	***	0.639	***	0.686	***
	(0.046)		(0.048)		(0.046)	
Social transfers to households	0.061	***	0.274	***	0.321	***
	(0.005)		(0.021)		(0.025)	
Government investment	0.432	***	0.569	***	0.561	***
	(0.042)		(0.053)		(0.048)	
Production grants	0.010	***	0.039	***	0.064	***
	(0.001)		(0.004)		(0.006)	
Households' labor income tax	0.065	***	0.300	***	0.375	***
	(0.006)		(0.024)		(0.032)	
Corporate income tax	0.014	***	0.043	***	0.071	***
	(0.001)		(0.001)		(0.002)	
Social security contributions	0.088	***	0.292	***	0.413	***
	(0.007)		(0.014)		(0.027)	
Consumption tax	0.040	***	0.172	***	0.271	***
	(0.003)		(0.020)		(0.037)	
Regional tax on economic activities	0.005	***	0.051	***	0.083	***
	(0.002)		(0.004)		(0.007)	
Excise duty on energy products	0.006	***	0.035	***	0.085	***
	(0.001)		(0.001)		(0.010)	

(1) Standard errors in parenthesis. ***, **, *denote significance at 1%, 5% and 10% levels, respectively.

Furthermore, the tax multipliers estimated in each of the sample periods are very close (especially for shocks to indirect tax rates), thereby confirming the near absence of regime dependence. Finally, an exception is represented by the output responses to changes in the social security contribution rate, whose magnitude is estimated to decline in both samples compared with the values given in Table 1.

The statistical framework of the stochastic simulation is extended to incorporate the case of dynamic simulations. The out-of-sample simulation period is increased to five steps ahead (so that each repetition consists in one dynamic simulation over the period of interest) and allows us to compute the interim (cumulated) fiscal multipliers for both time periods (Table 5 and Table 6 display up to 3-year interim multipliers). The estimates show that, irrespective of the specific estimation sample, the cumulative spending multipliers peak after about 3 years. The multipliers for government investment and consumption are significantly larger than those associated with other fiscal variables. Multipliers are small for general transfers, especially in terms of the impact. They increase considerably over time but remaining below those for government spending. Smaller multipliers are estimated for consumption taxes and labor income taxes, and even smaller multipliers are found for corporate income taxes. According to our findings, the path of tax multipliers is estimated to peak over a longer timespan (after 3 or 4 years), with the exception of the output responses to changes in excise duties and social security contributions. Overall, these dynamics are consistent with the findings prevailing in the empirical literature (among others, Coenen et al. 2012).

A statistical test of the significance of differences between each fiscal multiplier estimated in the two sample periods is also carried out. This assessment is performed on the basis of the stochastic simulation of the model, using the expected values computed from the *J* repetitions and the estimated standard errors. As shown in Table 7, the size of differences between the spending multipliers computed, respectively, for the full sample and the pre-crisis period are found to be statistically significant (at the 1 percent significance level) except for production grants, providing empirical support to the assumption of larger fiscal multipliers over the full-sample period. This evidence holds for both the impact multipliers and the interim multipliers. With regard to the tax multipliers, the differences in the size of impact multipliers are statistically significant for household income tax , corporate income taxes and consumption tax. Conversely, they are not significant in the other cases (social security contributions, regional tax on economic activity, excise duty on energy products). However, these deviations become statistically different from zero as the time horizon lengthens, with the sole exception of social security contributions.

The larger size of full-sample fiscal multipliers compared with the pre-crisis multipliers entails changes in model parameters, but to an extent that does not imply structural instabilities. However, given the linearity of the model, parameter estimates are independent of the state of the economy and might fail somewhat to properly account for the development of some specific variables, especially when they move close to the extremes of their range of variation. This could be the case of the "depressive" behavior of some variables (nominal interest rates constrained at the zero lower bound and low core inflation) observed during the crisis period that began in 2008.[15]

[15] Such special conditions could lead to an overestimation of the full-sample fiscal multipliers, as they could be affected by potential nonlinear effects. The overestimation of fiscal multipliers, depending on the structural characteristics of the models, is a debated issue in the empirical literature. As an example, the estimates reported in Auerbach and Gorodnichenko (2012a) are obtained using a nonlinear model under the assumption of "absorbing states", i.e., that an expansionary fiscal spending shock is unable to draw the economy out of a recession. Ramey and Zubairy (2018) demonstrate that this assumption may lead to an overestimation of spending multipliers. Using nonlinear techniques but relaxing the absorbing-state assumption they find that fiscal spending multipliers in recessions are, on average, not larger than those in good times. The issue of the overestimation of fiscal multipliers is also dealt with in Section 5 of this paper.

Table 7 – Comparing fiscal multipliers (1)

Fiscal variables	1-year		2-year		3-year	
Intermediate consumption	0.341	***	0.308	***	0.301	***
	(0.072)		(0.086)		(0.071)	
Social transfers to households	0.118	***	0.223	***	0.233	***
	(0.013)		(0.036)		(0.036)	
Government investment	0.504	***	0.482	***	0.529	***
	(0.080)		(0.106)		(0.125)	
Production grants	0.010		0.014		0.046	
	(0.163)		(0.063)		(0.045)	
Households' labor income tax	0.128	***	0.254	***	0.279	***
	(0.012)		(0.036)		(0.042)	
Corporate income tax	0.007	***	0.012	***	0.011	***
	(0.001)		(0.002)		(0.005)	
Social security contributions	0.011		0.031		0.011	
	(0.025)		(0.037)		(0.050)	
Consumption tax	0.047	***	0.135	***	0.124	***
	(0.007)		(0.032)		(0.045)	
Regional tax on economic activities	0.029		0.064	**	0.059	***
	(0.094)		(0.032)		(0.021)	
Excise duty on energy products	0.029		0.128	***	0.185	***
	(0.058)		(0.024)		(0.031)	

(1) Standard errors in parenthesis. ***, **, *denote significance at 1%, 5% and 10% levels, respectively.

Overall, the above findings provide empirical support for the qualitative assessment discussed in Section 3. First, the difference between the output responses to the same fiscal shock in the two samples is estimated to be even larger for the expenditure categories, as a result of the increase – compared with the version of the model not amended for structural breaks – in the size of fiscal multipliers estimated over the full-sample and, conversely, their reduction when only the information prior to the outbreak of the financial crisis is accounted for. Second, the results on the statistical significance of the gap in the dimension of fiscal multipliers provide reliable evidence of the heterogeneity of output responses to the same fiscal shock and, mostly important for our aims, point to the evidence of higher multipliers in the recession period (from 2008 onwards). Providing estimates of fiscal multipliers for this period is the goal of the next section.

5 Deducing fiscal multipliers in the crisis period 2008-2014

This section is devoted to describing the methodology adopted to estimate the fiscal multipliers over the crisis period 2008-2014 (denoted M2). To this end, we develop a statistical methodology that allows us to settle the problem posed by the impossibility of obtaining a direct estimation of M2 multipliers. The estimate of M2 fiscal multipliers is not feasible for MeMo-IT, since the length of the crisis period is too short in relation to the high dimensionality of the macro-econometric model. To circumvent this problem, we follow a strategy designed to perform an indirect estimate of M2 fiscal multipliers. It builds on two different estimates of the econometric model that are performed, respectively, using both the whole information set (covering the crisis period) and the data available up to the outbreak of the financial crisis. This procedure raises issues related to both the accuracy and efficiency of the estimates, which are reflected in the higher uncertainty of the statistical inference.

The methodological framework is developed starting from the seminal contribution of Chow (1960) and the theoretical literature which followed (Fisher, 1970; Ali and Silver, 1975; Jayatissa, 1977; Watt, 1979; Othani and Kobayashi, 1985). Based on these contributions, we find that, under the assumption of no structural breaks and homoskedasticity, the regression parameters estimated on the whole sample can be represented as the weighted average of the parameters estimated over, say, two subsamples, where the weights are obtained as the ratio of the inverse of variance-covariance matrices. In the second part of the section, we build on this framework to obtain the estimates of the fiscal multipliers for the crisis period.

5.1 Methodology

The starting point of this analysis is the final-form representation of the system introduced in Section 3. For ease of presentation, the specification of the final form is assumed with no lags in the exogenous variables, so to obtain a static representation of the model. This restriction does not limit the generality of the results, which can be extended to the case in which $\boldsymbol{\Pi}(L)$ is a polynomial in the lag operator. In matrix form, the system can be denoted as

$$y_t = \boldsymbol{\Pi} x_t + u_t$$

The stochastic assumptions made for the structural form have direct implications for the stochastic disturbance terms of the reduced form: u_t are stochastically independent with respect to x_t and uncorrelated over the sample; the covariance matrix of u_t is assumed constant over the sample; the sample data provides finite sample moments. Under these assumptions, the conditions of both the Gauss-Markov theorem and the Least Squares Consistency Theorem are satisfied for the final-form equations, so that $\boldsymbol{\Pi}$ can be estimated using the least squares estimator, which is a consistent although inefficient estimator of the equation-by-equation regression of $\boldsymbol{Y}$ on $\boldsymbol{X}$ (Hsiao, 1997). In what follows, the focus is on the *j*-th equation of the final-form representation of the model,

$$y_{t,j} = \sum_{i=1}^{k} \pi_{ij} x_{it,j} + u_{t,j}. \tag{7}$$

The objective is to derive an expression for the parameters π_{ij} (i=1,...,k) in terms of the coefficients of the same model (7) but estimated on partitions of the whole sample. To this aim, we adopt the theoretical framework developed in Chow (1960) under the stochastic assumptions of homoskedasticity and serial incorrelation of residuals.[16]

Assume that the sample t is split into two non-overlapping subsamples of size n_1 and n_2, where $n_1 > k$ and $n_2 > k$, so that equation (7) can be denoted as (in compact form and removing the subscript j)

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{8}$$

where subscript 1 denotes the first sample t_1 of n_1 observations, subscript 2 refers to the second sample t_2 of n_2 observations, X_1 and X_2 are nonsingular ($n_1 \times k$) and ($n_2 \times k$) matrices, and π_1 and π_2 are the vectors of the k regression coefficients. This representation implies that the parameters in the two subsamples are different ($\pi_1 \neq \pi_2$). The least squares estimators of π_1 and π_2, denoted by $\hat{\pi}_1$ and $\hat{\pi}_2$, are

$$\begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \end{pmatrix} = \begin{pmatrix} X_1{}'X_1 & 0 \\ 0 & X_2{}'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1{}' & 0 \\ 0 & X_2{}' \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} + \begin{pmatrix} (X_1{}'X_1)^{-1}X_1{}' \\ (X_2{}'X_2)^{-1}X_2{}' \end{pmatrix} \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix} \tag{9}$$

which implies

$$\begin{pmatrix} X_1{}'X_1 & 0 \\ 0 & X_2{}'X_2 \end{pmatrix} \begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \end{pmatrix} = \begin{pmatrix} X_1{}' & 0 \\ 0 & X_2{}' \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1{}'y_1 \\ X_2{}'y_2 \end{pmatrix}. \tag{10}$$

The variance-covariance matrix of parameter estimates is

$$\Omega_a = \sigma_1^2 (X_1{}'X_1)^{-1} + \sigma_2^2 (X_2{}'X_2)^{-1} = \Omega_{1,a} + \Omega_{2,a} \tag{11}$$

where $\Omega_{1,a}$ is the covariance matrix of the parameter estimates from the first n_1 observations, $\Omega_{2,a}$ is the covariance matrix of the estimates from the subsample t_2, σ_1^2 and σ_2^2 are the variances of residuals of the regressions on each subset t_1 and t_2, respectively.

Under the assumption of equality of the parameters π_1 and π_2 , ($\pi_1 = \pi_2 = \pi$), equation (8) reduces to

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \pi + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{12}$$

and the least square estimator of π, denoted by $\hat{\pi}$, can be written as

[16] Toyoda (1974) and Schmidt and Sickles (1977) show that the Chow test is not robust to heteroscedasticity. Gupta (1978, 1982), and implicitly Zellner (1962) and Theil (1971) develop an alternative test based on a likelihood ratio that deals with the problem. Jayatissa (1977) also introduces a test for structural change under heteroscedasticity.

$$\left[(X_1'X_2')\begin{pmatrix}X_1\\X_2\end{pmatrix}\right]\hat{\pi} = (X_1'X_2')\begin{pmatrix}y_1\\y_2\end{pmatrix} = X_1{}'y_1 + X_2{}'y_2 \tag{13}$$

with variance-covariance matrix

$$\Omega_0 = \sigma^2(X_1'X_1 + X_2'X_2)^{-1} \tag{14}$$

where σ^2 is the variance of residuals of the OLS estimates in (13). Taking account of (10) and (13) (Chow 1960), it follows that

$$(X_1'X_1 + X_2'X_2)\hat{\pi} = X_1'X_1\hat{\pi}_1 + X_2'X_2\hat{\pi}_2 \tag{15}$$

which implies

$$\hat{\pi} = X_1'X_1(X_1'X_1 + X_2'X_2)^{-1}\hat{\pi}_1 + X_2'X_2(X_1'X_1 + X_2'X_2)^{-1}\hat{\pi}_2 \tag{16}$$

Assuming equality of coefficients, expression (16) states that the parameter *k*-vector $\hat{\pi}$ estimated over the whole sample can be represented as a linear combination of the OLS coefficient vectors, $\hat{\pi}_1$ and $\hat{\pi}_2$, estimated in each subsample. The inner product $X_s'X_s$ (*s*=1, 2) is the contribution of the estimated parameters in subsample t_s to the estimates of $\hat{\pi}$. Under the homoskedasticity assumption, ($\sigma_1^2 = \sigma_2^2 = \sigma^2$), premultiplying both sides of (15) by σ^2, we obtain

$$\hat{\pi} = (\Omega_{1,a})^{-1}\Omega_0\hat{\pi}_1 + (\Omega_{1,b})^{-1}\Omega_0\hat{\pi}_2 \tag{17}$$

According to equation (17), the larger the information content, the lower the variance of the parameter estimates in subsample t_s, and the higher the weights attached to the coefficients. It follows that the influence of precise estimates is large, while the influence of less precise estimates plays a smaller role. This is because the reciprocal of the variance has exactly the same property.

In the simplifying case where $\hat{\pi}$ is a (2×1) random vector (*k*=2), expression (17) can be represented as

$$\begin{pmatrix}\hat{\pi}_{1,t}\\\hat{\pi}_{2,t}\end{pmatrix} = \begin{pmatrix}\gamma_{1,t1} & \\ \delta_{21,t1} & \gamma_{2,t1}\end{pmatrix}\begin{pmatrix}\hat{\pi}_{1,t1}\\\hat{\pi}_{2,t1}\end{pmatrix} + \begin{pmatrix}\gamma_{1,t2} & \\ \delta_{21,t2} & \gamma_{2,t2}\end{pmatrix}\begin{pmatrix}\hat{\pi}_{1,t2}\\\hat{\pi}_{2,t2}\end{pmatrix}$$

where $\gamma_{1,ts}$ ($\gamma_{2,ts}$) is the ratio of the reciprocal of variance for $\hat{\pi}_1$ ($\hat{\pi}_2$) in subsample t_s (*s*=1, 2) to the variance of the estimates on the whole sample; $\delta_{1,ts}$ ($\delta_{2,ts}$) is the same ratio but in terms of covariances.

In the following, we assume that the contribution of the covariance ratios is negligible for the estimation of $\hat{\pi}_{i,t}$, and the set of covariance weights in both subsamples is restricted to zero: $\{\delta_{il}\}_{ts} = 0$ (*i*≠*l*, *i*,*l*=1,2, *s*=1,2). This implies the assumption that fiscal shocks are orthogonal to each other (see Romer and Romer, 2010, for a similar assumption).[17] When the covariance terms are restricted to zero, the formulation for $\hat{\pi}_{i,t}$ (*i*=1,2) becomes

$$\hat{\pi}_{i,t} = \gamma_{i,t1}\hat{\pi}_{i,t1} + \gamma_{i,t2}\hat{\pi}_{i,t2} \tag{18}$$

The inverse variance weights are such as to minimize the variance of the weighted average estimator of *j* independent statistics (Hedges and Olkin, 1985). Given that $\hat{\pi}_{i,ts}$ are uncorrelated across subsamples with variance $V_{i,ts}$, the variance of the weighted mean estimator is $V_{i,t} = \sum_s \gamma_{i,ts}^2 V_{i,ts}$. To get the optimal weights $\gamma_{i,ts}$, the variance $V_{i,t}$ is minimized subject to the constraint that the sum of the weights is normalized to unity,

$$\mathbf{min}_{\gamma_{i,ts}} \mathrm{V}_{i,t} = \sum_{\mathrm{s}} \gamma_{\mathrm{i,ts}}^2 \mathrm{V}_{\mathrm{i,ts}} - \gamma_0 \left(\sum_{\mathrm{s}} \gamma_{\mathrm{i,ts}} - 1\right) \tag{19}$$

where γ_0 is the Lagrangian multiplier to enforce the constraint. Equating to zero the first derivative with respect to $\gamma_{\mathrm{i,ts}}$, we obtain $\gamma_{\mathrm{i,ts}} = (\gamma_0/2)/\mathrm{V}_{\mathrm{i,ts}}$, and taking sums on both sides, $2/\gamma_0 = 1/\sum \mathrm{V}_{\mathrm{i,ts}}$; thus, the optimal individual normalized weights are

$$\gamma_{i,ts} = \frac{1}{V_{i,ts}} \left(\frac{1}{\sum V_{i,ts}}\right)^{-1} = \frac{\omega_{i,ts}}{\sum_{\mathrm{s}} \omega_{i,ts}} \tag{20}$$

Substituting this definition into (19), the minimum variance of the weighted mean estimator is $V_{i,t} = 1/\sum V_{i,ts}$. The optimal individual weights, $\omega_{i,ts} = 1/V_{i,ts}$, are inversely proportional to the variance of the corresponding statistics, and the minimum variance unbiased estimator, among all possible unbiased weighted averages of, say, *s* independent statistics, being

$$\hat{\pi}_{i,t} = \frac{\sum_{\mathrm{s}} \omega_{i,ts} \widehat{\pi}_{i,ts}}{\sum_{\mathrm{s}} \omega_{i,ts}} = \gamma_{i,t1}\hat{\pi}_{i,t1} + \gamma_{i,t2}\hat{\pi}_{i,t2} \tag{21}$$

and its variance is

$$V_{i,t} = \gamma_{i,t1}^2 V_{i,t1} + \gamma_{i,t2}^2 V_{i,t2} + 2\gamma_{i,t1}\gamma_{i,t2}\rho\sqrt{V_{i,t1}V_{i,t2}} \tag{22}$$

where $V_{i,t1}$ and $V_{i,t2}$ are the variances of the *i*-th parameter in the respective subsample t_s (*s*=1,2). The parameter ρ allows us to account for the potential correlation between the estimates of the *i*-th fiscal multiplier across subsamples.

The inverse variance weights are optimal in the sense that the estimator of the population mean has minimum variance only if the variances are known. Since variance is usually unobserved, then estimates of variances are used, with the result that the estimator is no longer optimal, but it is unbiased (Keller and Holkin 2004).

[17] This approach might result in the over-estimation of fiscal multipliers. Favero and Giavazzi (2009) show that, under this representation, tax shocks are also orthogonal to any other macro shock (productivity shocks, shifts in government spending or in monetary policy) as well as to the lagged values of other excluded macro variables. Once these assumptions are relaxed, they find multipliers that are much smaller than the estimates provided by Romer and Romer (2010).

Expression (21) can be viewed as analogous to the methodology for the synthesis of summary estimates widely adopted in the framework known as meta-analysis (Hedges and Olkin 1985; Borenstein et al. 2009). Most of the statistical procedures in meta-analysis are usually based on the estimation of the average of summary statistics, with research into the best procedure for averaging being a point of remarkable methodological interest.

5.2 Procedure to estimate fiscal multipliers in the crisis period

The aim of this sub-section is to present the procedure adopted for the estimation of fiscal multipliers in the crisis period (2008-14). Given the above framework, the inference on the final-form parameter $\pi_{i,t2}$ can be carried out by reformulating expression (21),

$$\hat{\pi}_{i,t2} = \frac{(\hat{\pi}_{i,t} - \gamma_{i,t1}\hat{\pi}_{i,t1})}{\gamma_{i,t2}} \tag{23}$$

But this approach is not yet feasible, since not all the elements on the RHS of equation (23) are known. Some are estimated, namely the parameters pertaining to both the whole sample, $\hat{\pi}_{i,t}$, and the first subsample, $\hat{\pi}_{i,t1}$, along with the corresponding sample variances, $\hat{V}_{i,t}$ and $\hat{V}_{i,t1}$, respectively. The unknowns concern the weights $\gamma_{i,ts}$, which , according to (20), can be computed once $V_{i,t2}$ is obtained. As a result, to perform inference on $\pi_{i,t2}$, an estimate of the corresponding variance is required.

In the following, the variance of the *i*-th fiscal multiplier in the second subsample ($V_{i,t2}$) is estimated through an iterative procedure based on the minimization of a specific loss function attributed to Stein (James and Stein 1961) [18]

$$\mathbf{min}_{V_{i,t2}}(L(\hat{V}_{i,t}, \tilde{V}_{i,t})) = \frac{\tilde{V}_{i,t}}{\hat{V}_{i,t}} + \log\left(\frac{\tilde{V}_{i,t}}{\hat{V}_{i,t}}\right) - 1 \tag{24}$$

The minimization process is subject to the constraint that $V_{i,t2}$ must be positive, where $\hat{V}_{i,t}$ is the target value for the variance of $\hat{\pi}_{i,t}$ obtained from the stochastic simulation of the model, $\tilde{V}_{i,t}$ is the variance estimator as denoted by equation (22).[19] Once $V_{i,t2}$ is

[18] In this framework, the Stein loss (also called entropy loss or Kullback-Leibler loss function) has the desirable property that penalizes gross underestimation as gross overestimation since, for any fixed value of $\hat{V}_{i,t}$, $\mathrm{L}(\hat{V}_{i,t}, \tilde{V}_{i,t}) \rightarrow \infty$ as $\hat{V}_{i,t} \rightarrow 0$ or $\hat{V}_{i,t} \rightarrow \infty$. By contrast, the squared error loss is not appropriate for variance estimation because underestimation is penalized less (finite penalty) than overestimation (infinite penalty). For a survey on the Stein loss function, see Casella and Berger (2002), Maatta and Casella (1990), Kubokawa (1999); applications are in Tong and Wang (2007, 2012).

[19] It consists of a one-dimensional constrained optimization. The iterative procedure can be broadly outlined as follows: a) an initial guess for $V_{i,t2}$ is supplied as the starting value, which is generally of the same magnitude of overall variance, and an estimate for $\tilde{V}_1$ is obtained; b) at iteration $j \geq 1$, a guess for $\tilde{V}_j$ is obtained by updating $L()_{j-1}$. If $L()_j$is smaller than $L()_{j-1}$ by at least a predetermined amount, then the counter *j* is increased and the step is repeated. Otherwise, $\tilde{V}_{j-1}$ is taken as the estimator. The procedure is performed using the R package nlminb().

estimated ($\hat{V}_{i,t2}$), then the weights $\gamma_{i,ts}$ are computed ($\hat{\gamma}_{i,ts}$), and $\hat{\pi}_{i,t2}$ is obtained through the identity (23).

In order to apply this procedure for the estimation of $V_{i,t2}$, equation (22) must be exactly identified, so that the parameter ρ is set as exogenous. For the calibration of this parameter, we refer to the literature investigating the size of fiscal multipliers depending on the position of the business cycle (see also section 2). Some studies report the output responses to shocks to fiscal variables in both the expansionary and recessionary regimes over several time periods. That evidence is used to obtain indications of the degree of co-movement between the state-dependent multipliers. For Italy, Batini et al. (2012) find that the response of output following a spending (consolidation) shock exhibits a similar pattern both in expansion and in recession. As for tax multipliers, output reacts more asymmetrically in the short run, but the evolution in both regimes becomes more similar subsequently. As a result, the correlation between the fiscal multipliers over the business cycle is generally positive, being larger for the response of output to spending shocks, and smaller for tax shocks. Cimadomo and D'Agostino (2015) report a comparable result for the government spending multipliers.

However, it should be considered that the two subsamples used in this study do not properly identify distinct business-cycle regimes: the first subsample spans a long time period (from 1970 until 2007), thus including both recessions and expansions, while the second basically covers the period of the Great Recession, an unprecedented downturn compared with other phases of the Italian business cycle. We expect that the evolution of the *i*-th fiscal multiplier, specific to an exceptional recessionary period, is only weakly related to its development in the period before the crisis.

To evaluate developments in fiscal multipliers in recession compared with their evolution over a period that includes both peaks and troughs, we refer to Auerbach and Gorodichenko (2012b). They provide estimates of the impulse response functions to several fiscal shocks specific to the business cycles regimes for the US economy, along with the resulting impulse responses from a model with no regime shifts following the specification of Blanchard and Perotti (2002). Overall, the output response in recession compared with its development when not controlling for state dependence is generally symmetrical, although the output response is heterogeneous depending on the nature of the fiscal shock. The co-movement is large and positive for shocks to government purchases, it is milder in the case of investment spending, and close to zero or even slightly negative for the tax shocks. Caggiano et al. (2015) report similar results for the output reaction to a fiscal-news (anticipated) government spending shock.[20] In the light

[20] As regards the development of fiscal multipliers in recessions and expansions, Auerbach and Gorodichenko (2012b) find that the output responses to an increase in government expenditure diverge over time (Figure A4 for government purchases; Figure A5 for investment spending), implying a negative correlation between state-dependent multipliers; the same applies, albeit to a smaller extent, for tax shocks (Figure A1). By contrast, using the same model, Caggiano et al. (2015) report that, after a shock to government purchases, output reacts symmetrically in both the expansion and recession regimes, with the increase in expansion being milder than in recession, and subsequently the effect vanishes and output

of the above evidence, the range of values for parameter ρ can be restricted to take discrete values in the interval (0,1].

The estimation of the multipliers for the crisis period is obtained assuming that the covariance between the same fiscal multiplier across the two subsamples is zero. These estimates represent our baseline results. As a sensitivity check we also analyze cases with $\rho \neq 0$, i.e. taking account of the heterogeneous values that this parameter may assume depending on the nature of the fiscal shocks; this is done in sub-section 5.4 devoted to the robustness analysis. Setting $\rho = 0$, expression (22) simplifies to

$$V_{i,t} = \gamma_{i,t1}^{2} V_{i,t1} + \gamma_{i,t2}^{2} V_{i,t2}$$

where the variance in each of the two subsamples, $V_{i,t1}$ and $V_{i,t2}$, is not restricted to be the same. In the presence of structural changes, assuming that the variance of fiscal shocks does not change over time may be unrealistic (see Cimadomo and D'Agostino 2015; Fazzari et al. 2015), while it is likely to have increased compared with the values before the recent economic and sovereign debt crisis. The extent of such changes is differentiated depending on the nature of the fiscal instrument.[21]

5.3 Baseline results

The baseline multipliers are obtained from a procedure that involves two sets of fiscal multipliers: the full-sample multipliers and the pre-crisis multipliers. Both of them denote the real output responses to fiscal shocks of the same magnitude. As we evaluate the effects on real GDP of many different fiscal instruments, the fiscal impulses are normalized so that the size of the discretionary shock in each case represents a permanent increase in spending or a permanent decline in revenues equal to 1 percent of nominal GDP. Therefore, the baseline multipliers can be interpreted as the impact on output of normalized fiscal shocks in the crisis period. The impacts are expressed in percentage points. Results are reported in Table 8.

---

declines in both regimes (Figure 3). This different pattern is related to the fact that in the model of Caggiano et al. (2015), the economic system hit by the fiscal shock is allowed to switch from one state of the economy to another, while in Auerbach and Gorodichenko (2012b) the economy remains in the same state of the business cycle. That is, expansionary fiscal spending shocks are, by construction, not allowed to drive the economy out of a recession. This assumption provides an 'upper bound' for the estimates of the fiscal multiplier in recessions, because the effects of fiscal spending are not decreasing as the economy exits a recession (Caggiano et al. 2015).

[21] Cimadomo and D'Agostino (2015) state that, in presence of structural changes, assuming that the variance of shocks has not changed over time may be inappropriate; Fazzari et al. (2015) use a scale factor to allow the size of the shocks to change with the Great Moderation, but use a constant variance-covariance matrix.

Table 8 – Multipliers in the recession period – Baseline estimates (1)

Fiscal variables	Point estimates		
	1-year	2-year	3-year
Intermediate consumption	1.302	1.092	0.988
	[0.764, 1.840]	[0.167, 2.017]	[0.063, 1.913]
Social transfers to households	0.586	1.261	0.608
	[0.548, 0.623]	[1.111, 1.413]	[-0.095, 1.311]
Government investment	2.664	2.702	2.904
	[2.362, 2.966]	[2.321, 3.082]	[2.561, 3.247]
Production grants	-0.249	0.006	0.265
	[-0.723, 0.225]	[-0.020, 0.032]	[0.223, 0.307]
Households' labor income tax	0.634	0.824	0.653
	[0.593, 0.675]	[0.518, 1.131]	[-2.271, 3.578]
Corporate income tax	0.025	0.06	0.12
	[0.016, 0.035]	[0.031, 0.089]	[0.105, 0.135]
Social security contributions	0.136	0.428	0.462
	[0.083, 0.188]	[0.327, 0.529]	[0.267, 0.657]
Consumption tax	0.249	1.199	0.742
	[0.224, 0.274]	[1.094, 1.304]	[0.449, 1.034]
Regional tax on economic activities	0.134	0.336	0.344
	[-0.140, 0.407]	[0.304, 0.369]	[0.296, 0.392]
Excise duty on energy products	0.135	0.602	0.903
	[-0.034, 0.304]	[0.595, 0.609]	[0.835, 0.972]

(1) 95% confidence intervals in square brackets.

We find that all of the cumulative multipliers in the crisis period (2008-14) are larger than their pre-crisis values (1970-2007). The size of the differences between the multipliers in the two different periods varies depending on the nature of the fiscal instruments.

On the expenditure side, the results show that the cumulative fiscal multipliers for government investment and government consumption spending are the largest. The effect on output of changes in government investment provides the most effective impact on output, reaching a value of 2.6 in the first year and increasing to about 3 in the third year. With regard to the impact of a permanent shock in government purchases (intermediate consumption), the 1-year output multiplier is estimated at greater than unity (1.3), more than twice as much as the corresponding value estimated for the pre-crisis period. It starts to decay after the impact, approaching a value of about 1 in the third year. The change in social transfers to households has a considerably larger impact (0.6), about ten times its size in the pre-crisis period. It increases in the second year, becoming even larger than the second-year multiplier for government consumption spending. Overall, considering all the expenditure items, the spending multiplier in the crisis period is estimated on average at around 1 in the first year and 1.3 in the second year.

Also on the revenue side, fiscal multipliers in the crisis period, although lower than spending multipliers, increased compared with the pre-crisis period. The labor income tax multiplier is estimated at below 1 in the first year (0.6), but notably higher compared with its pre-crisis values; it increases in the second year (0.8) and then falls again subsequently. The impact on output of a shock to corporate income tax is particularly low (close to 0.1 in the three year period), although larger than the corresponding multipliers in the pre-crisis period. The multiplier for changes in social security contributions is larger (by about 1.5 times) compared with the values estimated before the crisis in the first and second year (0.14 and 0.43, respectively), while this difference narrows in the third period. The 1-year effect on output of a shock to consumption tax is estimated at 0.3, roughly 6 times the impact multiplier estimated in the pre-crisis period. The same proportion applies in the second year, when the multiplier quickly rises to 1.2. The fiscal multipliers for shocks to excise duties on energy products and also for changes in the regional tax on economic activities are estimated to be higher than to their pre-crisis levels. For example, the cumulative response to energy taxes reaches about 1 in the third year (it is about 0.1 for pre-crisis estimates). Overall, in the recession period, the average (cumulative) tax multiplier is around 0.3 in the first year and 0.6 in the second time period.

Although uncertainty is inevitably large in this kind of estimation, it is nonetheless possible to show that the size of the multipliers of the crisis period (2008-2014) is in general significantly larger than pre-crisis estimates (1970-2007), both on the expenditure and revenue sides. In terms of statistical significance, the majority of 1-year multipliers of the crisis period are significantly different from the corresponding pre-crisis estimates. This finding applies to all the expenditure multipliers and tax multipliers, while the null of no significance is accepted for the effects on GDP of changes in production grants and the regional tax on economic activities (Table 9). Larger differences compared with the pre-crisis values are found in the size of the fiscal multipliers on the expenditure side (similar results are in Auerbach and Gorodichenko, 2012b; Batini et al., 2012; Baum et al., 2012). We find that the difference with the pre-crisis values is greatest for changes in investment expenditure (2.2) and, to a smaller extent, intermediate consumption (0.8) and social transfers (0.5). Furthermore, as the period after the fiscal shock increases, the statistical significance of tax multipliers is generally retained, whereas some of them also gain statistical significance (such as the cumulative multipliers following a shock to regional taxes or excise duties on energy products). By contrast, some of the expenditure multipliers show a loss in significance. This is the case of 3-year multipliers for government spending and social transfers to households.

Table 9 – Baseline multipliers – Comparison with the fiscal multipliers in the pre-crisis period (1)

	$\widehat{\pi}_{t2} - \widehat{\pi}_{t1}$					
	1-year		2-year		3-year	
Intermediate consumption	0.780	**	0.454		0.302	
	(0.225)		(0.381)		(0.381)	
Social transfers to households	0.524	***	0.987	***	0.287	
	(0.016)		(0.065)		(0.288)	
Government investment	2.232	***	2.132	***	2.342	***
	(0.130)		(0.164)		(0.148)	
Production grants	-0.253		-0.033	**	0.202	***
	(0.194)		(0.011)		(0.018)	
Households' labor income tax	0.569	***	0.524	***	0.279	
	(0.018)		(0.127)		(1.196)	
Corporate income tax	0.011	**	0.017		0.048	***
	(0.004)		(0.012)		(0.007)	
Social security contributions	0.048	*	0.135	**	0.049	
	(0.023)		(0.043)		(0.084)	
Consumption tax	0.209	***	1.028	***	0.471	***
	(0.011)		(0.048)		(0.125)	
Regional tax on economic activities	0.128		0.286	***	0.261	***
	(0.112)		(0.014)		(0.021)	
Excise duty on energy products	0.128		0.567	***	0.819	***
	(0.069)		(0.003)		(0.029)	

(1) Standard errors in parenthesis. $\hat{\pi}_{t2} - \hat{\pi}_{t1}$ denotes the statistical test of the difference between the baseline multiplier and the corresponding multiplier in the pre-crisis period. It is performed through a two sample t-test assuming inequality of variances and $\rho = 0$. This statistic is distributed as Student's t with degrees of freedom given by the Satterthwaite's formula. ***, **, * denote significance at 1%, 5% and 10% levels, respectively.

5.3.1 Comparison with the literature

Overall, our baseline results seem broadly in line with the empirical evidence prevailing in the recent literature. We compare our findings on the output response to fiscal policy shocks during a recession with some of the existing estimates for Italy (where available) and for other countries and economic areas. These results are reported in Table 10 (for a more comprehensive discussion, see among others Mineshima et al. 2014).

As for the impact of a shock to government purchases, several authors find that, in the lower (recession) regime, the output multiplier increases well above 1. As for the Italian case, Locarno et al. (2013) present output multipliers assuming both a standard and an accommodative monetary policy stance during the fiscal stimulus. This latter case approximates a lower regime of the business cycle, where the output multiplier increases well above unity (for a five-year stimulus). Similar conclusions are in Warmedinger et al. (2015) and in Caprioli and Momigliano (2013), where the government consumption multiplier is greater than 2 (although it is not statistically significant compared with estimates for the higher regime). Kilponen et al. (2015) investigate how fiscal multipliers change depending on the fiscal rule (which serves to ensure fiscal sustainability) but not depending on the state of the business cycle. In response to a permanent reduction in government consumption, they find smaller 1-year multipliers (0.5 if the fiscal instrument is the household labor income tax , 0.7 if the fiscal rule is a lump-sum tax). As for other countries, Auerbach and Gorodnichenko (2012b) conclude that in the United States, during a slowdown the 4-quarter output multiplier is slightly below 1 (2.1 after 20 quarters). For the European Union, Coenen et al. (2012) estimate the average first-year multiplier to be equal to 0.9 and the average 2-year multiplier equal to 1.5 (assuming two years of monetary accommodation). Gechert et al. (2016), based on a meta-analysis of fiscal multipliers for several policy instruments and countries, point out that in the lower regime the impact on output of an increase in public expenditure is slightly below 2.

As for the government investment multiplier in the recession regime, to our knowledge there are no findings concerning the Italian case. Abiad et al. (2015) use a sample of 17 OECD economies and find that, in periods of low growth, a shock to public investment increases the level of output by about 1.5 percent in the first-year (3 percent in the medium term). Auerbach and Gorodnichenko (2012b) show that in the US the 4-quarter impact of investment spending is around 1.5 percent in the low regime (rising to 3.4 percent in the medium run). A similar result is in Gechert et al. (2015). As for the European Union, Rannenberg et al. (2015) simulate the euro area's fiscal consolidation between 2011 and 2013 by employing different macro models. They find that the 8-quarter cumulative multiplier ranges from 1.9 (NAWM) to 3.9 (QUEST III model). According to Coenen et al. (2012), the average 1-year instantaneous multiplier for the EU is equal to 1.5.

The literature also lacks findings for social transfers for Italy. The 2-year cumulated multiplier found in this paper (1.3) is roughly in line with the estimates reported in Rannenberg et al. (2015) for the EU: the 8-quarter GDP response to an increase in transfers ranges from 1 (QUEST III model) to 1.4 (NAWM). Similar results for the EU are in Coenen et al. (2012), while Gechert et al. (2016) estimate a considerably higher impact multiplier in the lower-regime (2.5). All these authors point out that the size of the response of the economy depends on the proportion of financially constrained households and on how well targeted the increase in transfers is during downturns.

Table 10 – Survey of literature on the size of fiscal multipliers during downturns

Fiscal shock	Source	Country/ Economic area	State of the economy (Monetary policy stance)	Fiscal multiplier
Government purchases	Locarno et al. (2013)	Italy	Lower regime (constant monetary policy rate)	1.37 (1-year), 1.13 (2-year)
	Warmedinger et al. (2015)	Italy	Recession	1.39 (1-year), 1.55 (2-year)
	Caprioli and Momigliano (2013)	Italy	Recession	>2 (4-quarters), >2.5 (8-quarters)
	Kilponen et al. (2015)	Italy	Average regime	<1 (depending on how the fiscal rule is specified).
	Auerbach and Gorodnichenko (2012b)	US	Recession	<1 (4-quarters), 2.1 (20 quarters)
	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	0.9 (average 1-year, EU); 1.5 (average 2-year; EU)
	Gechert et al. (2016)		Recession	<2 (on average)
Government investment	Abiad et al. (2015)	17 OECD economies	Recession	1.5 (1-year), 3 (medium term)
	Auerbach and Gorodnichenko (2012b)	US	Recession	1.5 (4 quarters), 3.4 (20 quarters)
	Gechert et al. (2016)		Recession	2 (on average)
	Rannenberg et al. (2015)	EU	Lower regime (increased share of credit constrained households)	1.9 (8 quarters, NAWM); 3.9 (8 quarters; QUEST III model)
	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	1.5 (average 1-year, EU); 1.6 (average 1-year, US)
Social transfers	Rannenberg et al. (2015)	EU	Lower regime (increased share of credit constrained households)	1 (8 quarters, QUEST III model), 1.4 (8 quarters, NAWM)
	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	>1 (average 1-year; EU);
	Gechert et al. (2016)		Recession	2.5 (1-year)
Government consumption and investment	Batini et al. (2012)	Italy	Recession	1.6 (1-year), 1.8 (2-year)
	Cimadomo and D'Agostino (2015)	Italy	Recession	1.5 (1980-90); 0.8/0.9 (until late 2000s); >1 (global crisis)
	Baum and Koester (2011)	G7 economies	Lower regime	1 (1-year; 2% GDP shock); 1.3 (1-year; 5% GDP shock)
	Caggiano et al. (2015)	US	Recession	3.1 (4-quarters), 3.4 (4-quarters, deep recession regime)
Households' labor income taxes	Locarno et al. (2013)	Italy	Lower regime (constant monetary policy rate)	0.4 (1-year), 0.6 (2-year), permanent stimulus
	Kilponen et al. (2015)	Italy	Average regime	0.2 (1-year), 0.4 (2-year)
	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	0.5 (average 1-year, EU), 0.2 (average 1-year, US)
Corporate income taxes	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	0.24 (average 1-year, US), 0.15 (average 1-year, EU)
	Mertens and Ravn, (2013)	US	Average regime	0.4 (1-quarter), 0.6 (1-year)
Consumption taxes	Locarno et al. (2013)	Italy	Lower regime (constant monetary policy rate)	0.16 (1-year), 0.23 (2-year) permanent stimulus
	Coenen et al. (2012)	EU, US	Lower regime (2-year monetary accommodation)	0.66 (average 1-year, EU) 0.61 (average 1-year, US)
	Rannenberg et al. (2015)	EU	Lower regime (increased share of credit constrained households)	0.5/0.2 (8 quarters, QUEST III model), 1/0.3 (8 quarters, NAWM)

When a broader definition of public expenditure is considered, our baseline estimates are slightly below the findings reported in the literature for Italy. Batini et al. (2012) evaluate the cumulative output multiplier for a shock to government expenditure (wages, current purchases of public goods and services and public investment) to be greater than 1 (in absolute value) and increasing over time during the recession period. Cimadomo and D'Agostino (2015) state that the macroeconomic effects of a shock to government spending (government consumption and investment) rose above unity during the global crisis (the average short-term multiplier was around 1). For the average of the G7 economies, Baum and Koester (2011) find that the 1-year cumulative multiplier to a positive spending shock in recession is 1 (when the shock represents an exogenous increase as large as 2 percent of GDP) or 1.3 (in the case of an exogenous increase as large as 5 percent of GDP). Caggiano et al. (2015) estimate (4-quarter) fiscal spending multipliers of more than 3 in the US during a recession. Similar results, considering a shock to both government consumption and investment, are obtained from models with some Keynesian features (Christiano et al. 2011; Eggertsson 2009; Woodford 2011).

With regard to the output response for shocks to average household labor income tax, our baseline results seem slightly on the high side. According to Locarno et al. (2013), for a permanent stimulus and assuming a constant interest rate, the 1-year fiscal multiplier is 0.4 (lower effects are found under standard monetary policy); the 1-year effect (in absolute value) falls to 0.2 in Kilponen et al. (2015), with the fiscal rule specified in terms of lump-sum taxes. Coenen et al. (2012) find that the 1-year instantaneous output multiplier to cuts in household labor income tax rates is 0.2 in the US and 0.5 in the EU (nearly invariant to the duration of monetary accommodation).

To our knowledge, there are no estimates for Italy of the corporate income tax multiplier in the recession regime, whose size seems mostly affected by country-specific factors (more than other tax instruments), thus increasing the heterogeneity of the estimates. Taking account of this, our estimates of the output response to shocks to the average corporate income tax are close to the findings reported in Coenen et al. (2012), but lower than those found in other studies (see for example Mertens and Ravn 2013).

As for the consumption tax multiplier, Locarno et al. (2013) find that, in the case of a permanent stimulus, the 1-year output multiplier for Italy is about 0.2 under an accommodative monetary policy regime (which does not change significantly assuming a standard monetary policy). Kilponen et al. (2015) present similar results, although not specific to a state of the economy. As for the EU, the first-year output multiplier is close to 0.7 (Coenen et al. 2012), while the 2-year cumulative fiscal multiplier ranges from 0.5 (QUEST III) to 1 (NAWM) (Rannenberg et al. 2015).

For the overall tax multiplier, our findings are in line with the results provided in Batini et al. (2012), especially for the 1-year multiplier. Auerbach and Gorodnichenko (2012b) estimate the US output response to a tax shock in recession at close to 0.4 after four quarters. A similar result is found in Baum and Koester (2011) for the average of the G7 countries; it is considerably lower in Gechert et al. (2016).

5.4 Sensitivity tests

As shown, the baseline estimates provide evidence of fiscal multipliers that are significantly higher during the crisis compared with the pre-crisis period. This result is in line with the majority of the empirical literature on the state-dependent effects of fiscal policy. The aim of this section is to discuss the robustness of the baseline estimates. Several sensitivity exercises are performed. First, with regard to the specification of the variance, as denoted in equation (22), the parameter ρ is allowed to take discrete values within the interval (0, 1]. Second, the fiscal multipliers in both subsamples are assumed to have the same population variance ($V_{t1} = V_{t2}$), therefore assuming that the variance does not change over time. Third, a different definition of the weights as reported in expressions (20) is used, approximating the reciprocal variance by the size of each subsample.

5.4.1 Correlation of the fiscal multipliers across subsamples

First of all, we allow for a variable degree of dependence of the multipliers across the subsamples. The correlation coefficient is calibrated to take discrete values within the interval (0, 1], thus accounting for a positive relation between the output response across the two subsamples to a given fiscal shock. As a result, the covariance of $\hat{\pi}_{i,t1}$ and $\hat{\pi}_{i,t2}$ will take on values satisfying the inequality $0 \leq Cov(\hat{\pi}_{i,t1}, \hat{\pi}_{i,t2}) \leq \gamma_{i,t1}\gamma_{i,t2}(V_{i,t1}V_{i,t2})^{1/2}$. The results for the 1-year output response to fiscal shocks are summarized in Table 11. As can be seen, the output multipliers for fiscal shocks tend to decrease as ρ rises: the closer ρ is to 1, the lower the magnitude of the impact multipliers. Few exceptions are represented by the output response to government investment shocks and to changes in consumption tax, where the size of the fiscal multipliers declines before increasing slightly as ρ approaches 1. When ρ=1, the output response to an intermediate consumption shock falls to 0.9 (from 1.3 for ρ=0), to 2.2 following a shock to government investment (including investment grants, from 2.4 for ρ=0), and to 0.4 for changes in social transfers to households (from 0.6 when ρ=0). Similar findings hold for the estimates of impact multipliers following changes in taxes. The output response falls in the case of a labor income tax shock (to 0.4 when ρ=0.9, from 0.6) and for changes in both the regional tax on economic activities and the excise duty on energy products (the magnitude decreases from 0.13 to 0.08 in both cases). The size of the impact multiplier declines to a lesser extent when shocks to consumption tax (0.2 when ρ=0.9) and social security contributions are considered, while it is almost unaffected following changes in corporate income taxes. All in all, when the correlation between the multipliers is accounted for, the magnitude of output responses generally decreases compared with the baseline estimates. But the confidence intervals overlap substantially with those obtained for ρ=0. This result suggests that the reaction of output to a fiscal shock is not significantly different even considering several degrees of correlation between the (same) fiscal multiplier across subsamples.

It is also worth noting that the estimated impact multipliers in the crisis period are, for higher values of ρ, still larger than the corresponding pre-crisis values. As a matter of fact, the difference between the output responses in the two sub-periods is statistically significant for the majority of fiscal instruments (Table 12). When shocks to intermediate consumption are considered, the difference in the output multipliers is statistically significant only for values of ρ marginally greater than zero (ρ<0.5). The multipliers for changes in production grants, social security contributions, the regional tax on economic activities and excise duty on energy products are slightly different but statistically equivalent to the pre-crisis multipliers.

Table 11 – Impact multipliers (1-year) in the crisis period for several values of the correlation coefficient (1)

	ρ =0.1	ρ =0.5	ρ =0.9
Intermediate consumption	1.134	0.972	0.946
	[0.441, 1.827]	[-0.490, 2.433]	[-1.417, 3.310]
Social transfers to households	0.519	0.406	0.367
	[0.479, 0.557]	[0.363, 0.448]	[0.327, 0.412]
Government investment	2.379	1.899	2.163
	[2.066, 2.691]	[1.556, 2.242]	[1.924, 2.403]
Production grants	-0.217	-0.162	-0.144
	[-0.707, 0.274]	[-0.701, 0.376]	[-0.710, 0.422]
Households' labor income tax	0.562	0.439	0.397
	[0.519, 0.604]	[0.393, 0.486]	[0.348, 0.446]
Corporate income tax	0.026	0.026	0.026
	[0.016, 0.036]	[0.016, 0.036]	[0.016, 0.036]
Social security contributions	0.130	0.119	0.116
	[0.075, 0.184]	[0.0597, 0.179]	[0.053, 0.178]
Consumption tax	0.222	0.177	0.194
	[0.196, 0.248]	[0.149, 0.206]	[0.173, 0.215]
Regional tax on economic activities	0.117	0.090	0.080
	[-0.167, 0.401]	[-0.222, 0.401]	[-0.247, 0.407]
Excise duty on energy products	0.118	0.091	0.081
	[-0.057, 0.293]	[-0.102, 0.283]	[-0.121, 0.283]

(1) 95% confidence intervals in square brackets.

Table 12 – Impact multipliers (1-year) in the crisis period for several values of the correlation coefficient – Comparison with the fiscal multipliers in the pre-crisis period (1)

	$\hat{\pi}_{t2}-\hat{\pi}_{t1}$							
	ρ =0		ρ =0.1		ρ =0.5		ρ =0.9	
Intermediate consumption	0.780	**	0.612	*	0.449		0.424	
	(0.225)		(0.289)		(0.610)		(0.987)	
Social transfers to households	0.524	***	0.457	***	0.345	***	0.306	***
	(0.016)		(0.017)		(0.019)		(0.021)	
Government investment	2.232	***	1.946	***	1.467	***	1.731	***
	(0.130)		(0.136)		(0.156)		(0.123)	
Production grants	-0.253		-0.221		-0.166		-0.148	
	(0.194)		(0.201)		(0.220)		(0.232)	
Households' labor income tax	0.569	***	0.496	***	0.374	***	0.332	***
	(0.018)		(0.019)		(0.021)		(0.023)	
Corporate income tax	0.011	**	0.012	**	0.012	**	0.012	**
	(0.004)		(0.004)		(0.004)		(0.004)	
Social security contributions	0.048	*	0.042		0.032		0.028	
	(0.023)		(0.024)		(0.027)		(0.030)	
Consumption tax	0.209	***	0.182	***	0.137	***	0.154	***
	(0.011)		(0.011)		(0.013)		(0.011)	
Regional tax on economic activities	0.128		0.112		0.084		0.075	
	(0.112)		(0.116)		(0.128)		(0.135)	
Excise duty on energy products	0.128		0.112		0.084		0.075	
	(0.069)		(0.072)		(0.079)		(0.083)	

(1) Standard errors in parenthesis. $\hat{\pi}_{t2}-\hat{\pi}_{t1}$ denotes the statistical test of the difference between the fiscal multiplier in the crisis period and the corresponding multiplier in the pre-crisis period. It is performed through a two sample *t*-test assuming inequality of variances. This statistic is distributed as Student's *t* with degrees of freedom given by the Satterthwaite's formula. The standard error of the difference accounts for the covariance between the multipliers when the correlation coefficient is assumed greater than zero. ***, **, * denote significance at 1%, 5% and 10% levels, respectively.

5.4.2 *Equality of variance in both subsamples ($V_{t1}=V_{t2}$)*

A second issue for the sensitivity analysis is to evaluate the effect on the estimates for the *i*-th multiplier $\hat{\pi}_{i,t2}$ when its variance in both subsamples is assumed to follow the same population variance, so that $V_{i,t1}=V_{i,t2}$, and observed differences are only due to sampling variability. Assuming $\rho=0$, the expression of the variance for the *i*-th full-sample multiplier becomes

$$V_{i,t}=V_{p,t}\left(\gamma_{i,t1}^{2}+\gamma_{i,t2}^{2}\right) \tag{25}$$

where

$$V_{p,t} = \frac{(n_1 - 1)\hat{V}_{i,t1} + (n_2 - 1)\hat{V}_{i,t2}}{n_1 + n_2 - 2}$$

is the pooled variance, which is a weighted average of the estimated variances in both subsamples, where the weights are ratios of the degrees of freedom. The reliability of this approach, although common in most parametric data analysis techniques, depends on the admissibility of the assumption of equal population variance. If this hypothesis is verified by the data, then the pooled variance provides the most accurate estimates of population variance.

In this framework, the estimation of the unobserved variance for the second subsample is obtained by substituting expression (25) for the definition of $\tilde{V}_{i,t}$ in the loss function to be minimized, as defined in (24). The assumption of equal population variance for the *i*-th multiplier and of no correlation of the output responses across subsamples implies that $V_{i,t2}$ is closer to the population variance. Indeed, the sample variances for the second subsample are estimated to be lower than the corresponding sample statistics obtained under the assumption of separate variance (the baseline case). This implies that the weights for the period t_2 are larger than those computed in the baseline exercise. In terms of (23), the estimated fiscal multipliers for the second period (referred as pooled multipliers, hereafter) are expected to be lower than the corresponding baseline estimates, $\hat{\pi}_{t2,pooled} < \hat{\pi}_{t2}$. This is confirmed by the results presented in Table 13 (column *a*). The most substantial reductions concern, among the estimates that are significantly different from zero, the output multipliers associated with changes in social transfers, investment expenditure, household labor income tax and consumption tax.

When we turn to a statistical difference, the results reported in Table 13 (column *a*) suggest that the pooled multipliers are statistically equivalent to the baseline multipliers, as the confidence intervals overlap for the majority of fiscal multipliers. Exceptions include the output responses for changes in household transfers, government investment, household labor income tax and consumption tax, which are significantly lower than the baseline estimates. But when the comparison is performed against the pre-crisis multipliers (Table 14, column *a*), the pooled estimates for the crisis period are significantly larger for most of the fiscal variable shocks: the null of no significant difference is accepted for changes in production grants, social security contributions, the regional tax on economic activities and excise duties on energy. Therefore, the baseline estimates can be considered to be bounded from below by the set of pooled multipliers.

Table 13 – Fiscal multipliers in the crisis period (1-year): pooled estimates and combined estimates (1)

	Pooled estimates $\widehat{\pi}_{t2,p}$ (*a*)	Combined estimates $\widehat{\pi}_{t2,c}$ (*b*)
Intermediate consumption	0.929 [0.533, 1.325]	2.716 [2.468, 2.964]
Social transfers to households	0.181 [0.095, 0.268]	0.822 [0.535, 1.111]
Government investment	0.968 [0.488, 1.448]	3.672 [2.776, 4.567]
Production grants	-0.054 [-1.168, 1.060]	-0.363 [-0.656, -0.071]
Households' labor income tax	0.198 [0.119, 0.276]	0.891 [0.605, 1.178]
Corporate income tax	0.021 [0.012, 0.030]	0.056 [0.027, 0.086]
Social security contributions	0.099 [-0.058, 0.255]	0.158 [0.065, 0.250]
Consumption tax	0.090 [0.049, 0.131]	0.343 [0.253, 0.433]
Regional tax on economic activities	0.035 [-0.608, 0.678]	0.191 [0.098, 0.284]
Excise duty on energy products	0.035 [-2.532, 2.602]	0.192 [0.100, 0.284]

(1) 95% confidence intervals in square brackets.

Table 14 – Fiscal multipliers in the crisis period (1-year): pooled estimates and combined estimates – Comparison with the fiscal multipliers in the pre-crisis period (1)

	Pooled estimates		Combined estimates	
	$\widehat{\pi}_{t2,p} - \widehat{\pi}_{t1}$		$\widehat{\pi}_{t2,c} - \widehat{\pi}_{t1}$	
	(*a*)		(*b*)	
Intermediate consumption	0.407	***	2.193	***
	(0.126)		(0.116)	
Social transfers to households	0.120	***	0.761	***
	(0.019)		(0.049)	
Government investment	0.535	***	3.239	***
	(0.127)		(0.179)	
Production grants	-0.058		-0.367	***
	(0.185)		(0.049)	
Households' labor income tax	0.132	***	0.826	***
	(0.019)		(0.049)	
Corporate income tax	0.007	***	0.042	
	(0.002)		(0.005)	
Social security contributions	0.011		0.070	***
	(0.031)		(0.023)	
Consumption tax	0.050	***	0.303	***
	(0.011)		(0.017)	
Regional tax on economic activities	0.029		0.186	***
	(0.107)		(0.016)	
Excise duty on energy products	0.029		0.186	***
	(0.426)		(0.015)	

(1) Standard errors in parenthesis. $\hat{\pi}_{t2} - \hat{\pi}_{t1}$ denotes the statistical test of the difference between the fiscal multiplier in the crisis period and the corresponding multiplier in the pre-crisis period. It is performed through an independent two sample *t*-test assuming equality of variances. ***, **, * denote significance at 1%, 5% and 10% levels, respectively.

5.4.3 Weighting by sample size

Where each subsample is expected to provide original information on the population mean that is not contained in the other, the sample mean in the two subgroups of data should differ substantially. The pooled estimates should be different, and likely lower, than the values obtained by first combining all the values together and then calculating the overall mean and variance. To see this, let us consider the subsamples $\{x_{t1}\}_{t1=1}^{n_1}$ and $\{x_{t2}\}_{t2=n_1+1}^{n_1+n_2}$, $(n_1 \neq n_2)$, and assume that the observations in each sample follow the same population variance. If we combine all of the data into a unique sample of size $n_1 + n_2$, this is equivalent to a collection of sample means with weights n_1 and n_2, respectively. In the more general case of m subsamples (s=1,…,m), the definition (21) of the weighted average with minimum variance becomes

$$\frac{\sum_{s=1}^{m} n_s\,\hat{\pi}_{i,ts}}{\sum_s\ n_s}$$

which is the weighted mean of m samples. As a further sensitivity analysis, we consider the application of this estimator, where the weights that minimize the variance are defined in terms of the sample sizes (referred to as the combined estimator, hereafter; see Hunter and Schmidt, 2004). Several studies have investigated the statistical performance of this weighting strategy compared with weighting in terms of the inverse variance (Hedges and Olkin 1985). Weighting by the inverse variance yields more accurate results, as the optimal weight is the inverse variance of each sample estimate. But, in the empirical applications, the population variance is usually unknown and must be estimated from the empirical data. Thus, the weights $\omega_{i,ts}$ are likely affected by sampling error and the estimator is no longer optimal, although it is unbiased (Keller and Holkin, 2004), yielding estimates that are less accurate with respect to the optimal estimator. Weighting by the sample size is, therefore, a reasonable alternative, as it closely approximates the inverse variance, and being less affected by sampling error, is fairly unbiased (Sanchez-Meca and Marin-Martinez 1998, 2010; Schmidt et al. 2009).

The empirical findings, reported in Table 13 (column *b*), show that the combined estimator yields larger impact multipliers compared with the baseline estimates. In terms of statistical significance, the combined estimates of the expenditure multipliers are significantly larger than the baseline multipliers, since the confidence intervals do not overlap (as in the case of intermediate consumption and government investment) or the overlap is only partial (social transfers). Conversely, all of the tax multipliers (with the exception of the output response to a labor income tax shock) are not statistically different from the corresponding baseline estimates, as the confidence intervals for the combined multipliers overlap with those of the baseline estimates. Furthermore, the statistical comparison with the pre-crisis multipliers (Table 14, column *b*) shows that the combined estimates are significantly different (and larger) for all of the fiscal variable shocks (except for corporate income tax). These results suggest that the 1-year baseline multipliers may be assumed to be bounded from above by the corresponding combined estimates.

All in all, we interpret the reported magnitudes of the multipliers to the several sensitivity checks as possible lower and upper bounds rather than as standard values. First, the pooled estimates are obtained under the assumption of common population variance (other than the mean) in both subsamples. But this hypothesis may only be weakly reliable, because the intensity and length of the downturn observed during the crisis were uncommon compared with previous recessions. Second, when the weighting is performed in terms of the sample size, the advantage is that the multipliers are obtained using an estimator that is theoretically unbiased, although they are on the upper side of the admissible results for the expenditure multipliers. As a possible motivation, weighting by sample size implies constant weights across the fiscal shocks, and so an assumption of homogeneous variance for all fiscal multipliers, although it seems more reasonable to assume that the fiscal policy instruments were affected differently during the crisis period, so that the variance of the shocks is heterogeneous across the fiscal variables. As a consequence, using constant weights could lead to output responses that are unreliable for changes in specific exogenous fiscal variables. We deem that more realistic estimates will fall between the extremes of the above sensitivity analysis, so we select the baseline multipliers as the reference estimates for the fiscal multipliers in the crisis period.

In Section 6 below we make use of these baseline estimates of the fiscal multipliers for the crisis period to evaluate to what extent their adoption, in substitution of the standard multipliers of the MeMo-It macroeconomic model, allows us to reduce the prediction error for the Italian GDP growth rate in the crisis period. Indeed, this traditional macroeconomic model, as well as almost all the other forecasting models for the Italian economy, significantly underestimated the extent of GDP contraction following the fiscal consolidation plan implemented in Italy at the end of 2011.

6 An application: impact of the fiscal consolidation plan of July-December 2011

The aim of this section is to evaluate the extent to which the use of multipliers specific to the crisis period reduces the forecast errors for the Italian GDP growth rate in the post-2011 recession, following the euro-area debt crisis and the fiscal consolidation adopted by the Italian government. The assessment is carried out in the form of a predictive validation exercise (see Draper et al., 1993). In the retrospective form of predictive validation, outcomes that have already been observed are predicted again using the baseline multipliers estimated in Section 5. Then, the observed outcomes are compared with their predictions.

In our framework, the observed outcome is represented by the official GDP growth rate for the years from 2012 to 2014, as reported in the Istat press release of March 2017. For the projection of GDP developments in the post-2011 recession, we refer to the Istat forecast exercise covering the period 2012-2013, as published in the economic outlook of May 2012.[22] This is the benchmark forecast: it allows us to estimate the growth forecast errors and, then, to assess the reduction in the prediction error once the baseline multipliers are used.

The Istat forecast exercise is taken into account for two main reasons. First, the Istat projection of May 2012 was carried out using MeMo-It macro-econometric model, which is currently used to perform the macroeconomic forecast of the Italian economy

[22] http://www4.istat.it/en/archive/economic+outlook/page/2.

two times a year. The fiscal multipliers underlying the Istat GDP projections, reported in Bacchini et al. (2013, 2015), are obtained from the estimates of structural parameters of the MeMo-It model over the 1970-2011 period.[23] Second, the Istat forecast for the years 2012-2013 is inclusive of the effects of the "Berlusconi-Monti manovra" (described in the Box 1), i.e. the consolidation measures adopted by the Italian governments in the second half of 2011 in reaction to the consequences of the sovereign debt crisis.

As for the timespan considered, we analyze developments in the 3-year period following the plan adoption. However, the Istat projection only covers the years 2012-2103, so that we have to extend the model-based outlook by one additional year. Specifically, we proceed by replicating the Istat forecast outturn as of May 2012 (using the version of the MeMo-It model provided by Istat in 2015). This is done by incorporating: *a*) the published Istat assumptions on the evolution of the international exogenous variables for the period 2012-2013 and, *b*) the policy measures of the "Berlusconi-Monti" fiscal consolidation plan, including the information concerning how the discretionary fiscal policy measures are specified in the Istat model. The extension of the forecast by an additional year to 2014 is implemented by taking account of the forecast of the exogenous variables available at the time the Istat exercise was carried out (April-May 2012), as well as of the specification of the planned budgetary measures for 2014.

Before examining the exercise, it has to be noted that large and persistent forecast errors were common among practically all the forecasters in the estimates they made for the Italian economy in the aftermath of the sovereign debt crisis and the consolidation plan. Essentially, the forecast errors underlying the Istat economic outlook of May 2012, which are the focus of this section, were not an exception. In fact, to understate the extent of GDP contraction was rather the rule, being largely in line with the prediction errors made by almost all the forecast exercises carried out at that time (see Box 2 for more on this).

[23] The fiscal multipliers are obtained from the version of the MeMo-It model provided by Istat in 2015. It is substantially similar to the model used by Istat to perform the forecasting exercise of May 2012. The estimated fiscal multipliers, although slightly different, are consistent with the figures reported in Bacchini et al. (2013, 2015).

Box 1 – The 2011 fiscal consolidation plan

The fiscal consolidation plan for 2012-2014 took shape in the second half of 2011 through three subsequent sets of measures that gradually increased the size of the adjustment as the contagion of the sovereign-debt crisis progressively involved the Italian economy. The Berlusconi government adopted the first two sets of measures in the summer of 2011 (July and August), while the third was decided in December by the Monti government (so called "Salvaitalia" decree). Overall, the fiscal consolidation plan amounted to a cut in the budget deficit, with respect to the unchanged policy scenario, of €81 billion in 2014, corresponding to a fiscal adjustment of 4.8 percent of nominal GDP (Table B1.1).

Table B1.1 – The 2011 overall fiscal consolidation plan (1)
(in millions of euros, unless stated otherwise)

	2011	2012	2013	2014
Expenditures				
Compensation of employees	0	0	-70	-1,440
Intermediate consumption	-896	-6,814	-10,602	-12,157
Production grants	352	4,481	1,531	1,431
Social transfers	53	-3,243	-8,454	-10,501
Other current expenditures	0	-2,022	-3,251	-3,301
Government investments	-280	-2,619	-2,177	-1,657
Capital transfers	534	1,543	-297	-157
Total expenditures	**-237**	**-8,674**	**-23,250**	**-26,342**
Revenues				
Direct taxes				
Households' labor income tax (IRPEF)	-129	2,291	10,258	10,023
Corporate income tax (IRES)	544	1,401	1,335	-132
Regional tax on econ. activities (IRAP)	0	863	-504	803
Other personal taxes	23	66	66	66
Other taxes on corporate revenue	74	94	997	1,983
Taxes on financial revenues	0	1,421	1,534	1,915
IMU and rents revaluation	0	10,660	10,930	11,330
Indirect taxes				
Excise duty on energy products	0	8,153	8,397	9,161
VAT and other indirect taxes	2,086	12,632	16,832	17,940
Social security contributions	-3	1,066	1,471	1,886
Other current revenues	13	138	-460	-881
Capital taxes	0	1,461	1,987	559
Total revenues	**2,609**	**40,244**	**52,841**	**54,652**
Net borrowing	2,846	48,918	76,091	80,994
% of GDP	*0.2*	*3.0*	*4.6*	*4.8*

(1) Difference with respect to the unchanged policy scenario.

The bulk of the consolidation measures were concentrated in 2012, reducing the budget deficit with respect to the trend scenario by a little less than €50 billion, an amount as large as 3 percent of GDP. An additional correction was planned for 2013, with a further cut in the deficit of €28 billion, corresponding to 1.6 percent of GDP, while only a marginal additional adjustment concerned 2014, with a further reduction in the deficit of about €5 billion (0.2 percent of GDP). As for the contribution of the different tranches of measures to the adjustment plan, the Monti decree of December 2011 essentially frontloaded the fiscal consolidation process with respect to the decrees of the Berlusconi government. Indeed, the "Salvaitalia" decree significantly increased the reduction of the 2012 deficit, adding an extra adjustment of about €20 billion and increasing the deficit cut in that year from 1.8 to 3 percent of GDP.

Table B1.1 shows the composition of the overall fiscal consolidation plan adopted in 2011. The adjustment mainly centered on the revenue side in 2012 (when more than 80 percent of the adjustment was due to revenue increases), while in 2013 it partially shifted towards the expenditure side (about 55 percent of the deficit squeeze in that year was due to spending cuts). Indirect taxes mostly contributed to the revenue increase in 2012, while direct taxes sustained the revenue increase in the following years. As for expenditures, larger cuts were planned throughout the 3-year period for intermediate consumption and payments for social transfers. Measures aiming at cutting public investment, together with those concerning the reduction of the other current expenditures, were concentrated in 2012.

In our exercise we seek to assess the reduction in the growth forecast errors when, leaving all other things equal, two factors are taken into account: the actual observed path of the exogenous variables and the fiscal multipliers for the crisis period that we estimated in Section 5 (baseline multipliers). As for the first factor, empirical findings show that the revision of the expected path of the exogenous variables concerning the international environment explains a considerable share of the forecast errors. For the Italian case, Busetti and Cova (2013) perform a counterfactual analysis and find that the effects of deterioration in the international scenario deducted almost 2 percentage points of GDP growth in the period 2012-2013. Felici et al. (2017) conduct a similar study and conclude that the sizeable forecast errors primarily reflect the ex-ante assumptions about developments in the exogenous variables over the post-2007 period. Regarding the second issue, Blanchard and Leigh (2013) demonstrated the key role of the underestimation of fiscal multipliers early in the crisis in causing large forecast errors.

In what follows, we implement an approach to incorporate output multipliers in macroeconomic projections that is widely used in applied macroeconomics for its ease of implementation (Batini et al. 2014b). Once the set of estimated fiscal multipliers is selected, and the fiscal shocks are estimated, the macroeconomic effects of the fiscal measures, as well as the lagged effects of past measures, are obtained by applying the fiscal multipliers to the planned discretionary fiscal measures. For a 3-year period, the cumulated effect of a specific fiscal shock in t_3 can be obtained as follows,

$$Cumulated\ effect = \hat{\pi}_1 \sum_{i=1}^{3} \widehat{FS}_{ti} + (\hat{\pi}_2 - \hat{\pi}_1) \sum_{i=2}^{3} \widehat{FS}_{t(i-1)} + (\hat{\pi}_3 - \hat{\pi}_2) \widehat{FS}_{t1} \qquad (26)$$

where $\hat{\pi}$ denotes the estimate of the cumulated fiscal multiplier, $\widehat{FS}$ is the discretionary change in the corresponding fiscal measure, t_i (i=1, 2, 3) represents the time of the fiscal shock so that, for i=1 , $\hat{\pi}_1$ is the 1-year impact multiplier, $\widehat{FS}_{t1}$ is the fiscal shock occurring in t_1 for the first time, and $\hat{\pi}_1 \widehat{FS}_{t1}$ is the first-year estimated output response to that fiscal shock. The expression for the cumulated effect of changes in a policy measure accounts for the overlapping effects of the past fiscal shocks due to the persistence multiplier effects.

Box 2 – The growth forecasts of the Italian economy after the fiscal consolidation plan

The growth forecasts for the Italian economy elaborated by different institutions after the emergence of the sovereign debt crisis and the subsequent adoption of the fiscal consolidation plan were characterized by multiple overestimation errors. Table B2.1 shows the GDP forecasts for the period 2012-2014 estimated by a panel of institutes during the 7 months following the adoption of the plan (the median values and the upper-lower bounds of these forecasts are also shown). Forecasts are grouped by the period in which they were produced, distinguishing three different phases: Winter (December 2011-January 2012), Spring (March-May 2012), Summer (June-July 2012). The table also shows GDP outturns as reported in the subsequent Istat releases (the March releases are considered).

Forecast errors were quite large. Even considering the huge uncertainty measured by the dispersion of the estimates (the distance between the minimum and maximum values of the forecast is ample, particularly for 2012 and 2013), the failure in predicting the depth and length of the recession was generally widespread across the forecasters. For 2012, GDP overestimation was significant in the winter and spring forecasts. The summer forecasts were a bit closer to the Istat preliminary figures for actual performance. The latter however would later be revised in subsequent Istat releases (showing a more severe recession), on the basis of new and more complete information. For 2013 and 2014, forecasting errors were even more persistent. In the 2012 summer estimates, forecasters (even the most pessimistic) still predicted that the recession would gradually dissipate during 2013 and a recovery would materialize in 2014.

Table B2.1 – GDP growth forecasts and outturns for the period 2012-2014
(percent change)

	2012	2013	2014
	GDP forecasts, median values; min-max values in parenthesis (1)		
Winter forecasts: December 2011-January 2012	-1.6 (-2.2; -0.4)	0.1 (-0.6; 0.6)	1.3 (1.0; 1.5)
Spring forecasts: March-May 2012	-1.5 (-1.9;-1.2)	0.3 (-0.4; 0.5)	1.0 (0.5; 1.2)
Summer forecasts: June-July 2012	-2.0 (-2.4;-1.9)	-0.3 (-0.4; 0.1)	1.0 (0.9; 1.1)
	GDP outturns in subsequent Istat releases		
Istat, National accounts release March 2013	-2.4		
Istat, National accounts release March 2014	-2.4	-1.9	
Istat, National accounts release March 2015	-2.8	-1.7	-0.4
Istat, National accounts release March 2016	-2.8	-1.7	-0.3
Istat, National accounts release March 2017	-2.8	-1.7	0.1

(1) Forecasters considered are Ministry of Treasury (December 2011 and April 2012), Bank of Italy (January and July 2012), Istat (April 2012), Ref (January, April and July 2012), Prometeia (January, April and July 2012), CER (March and July 2012), Confindustria (December 2011 and June 2012), IMF (January, April and July 2012), OECD (May 2012), EC (April 2012).

The overall impact of fiscal shocks in year t_2 consists of two effects: the 1-year impact of the fiscal shock occurring in the second year, $\hat{\pi}_1\widehat{FS}_{t2}$, and the 2-year impact pertaining to the change in the policy measure taking place in the first-year, $(\hat{\pi}_2-\hat{\pi}_1)\widehat{FS}_{t1}$; the difference between cumulated output multipliers $(\hat{\pi}_2-\hat{\pi}_1)$ provides the 2-year impact multiplier (needed to compute the second-year output response to the fiscal shock occurring in the previous period). As a result, for *i*=3, the cumulated effect on real output in t_3 is inclusive of the impact of the simultaneous fiscal shock, $\hat{\pi}_1\widehat{FS}_{t3}$, and of the lagged effects of the measures adopted in the previous years, $(\hat{\pi}_2 - \hat{\pi}_1)\widehat{FS}_{t2}$ and $(\hat{\pi}_3 - \hat{\pi}_2)\widehat{FS}_{t1}$.

The above scheme represents the reference framework for the estimation of the impact of discretionary changes in fiscal measures on the GDP growth rate. Two sets of fiscal multipliers are used: the first refers to the multiplier estimates consistent with Istat GDP projection (Tab. 1); the second concerns the fiscal multipliers specific to the crisis period (baseline multipliers, Tab. 8). The fiscal shocks are evaluated on the basis of the estimates of the fiscal policy measures reported in the budget documents and described in Box 1.

The same framework is also used to evaluate the impact on GDP growth rate of the revised path of the global exogenous variables. This is performed by replacing $\widehat{FS}$ with the deviation between the actual and the expected path of a given exogenous variable, where $\hat{\pi}_i$ (i=1, 2, 3) is the corresponding cumulated multiplier implicit in the Istat forecast. For each year of the forecast horizon, the estimated effect represents the variation in the growth forecast errors attributable to changes in the development of the exogenous variables. When this is added to the Istat GDP forecast of May 2012, all other things being equal, the output projection is adjusted for the discrepancies between the actual and the expected path of the exogenous variables.

In order to identify the contribution to the growth forecast errors of the underestimation of fiscal multipliers, a GDP projection under the assumption of a "no-policy change" scenario is needed (baseline GDP). It represents, all things being equal, the output developments that would have been observed if the policy measures had not been adopted. For each year of the forecast horizon, the baseline GDP growth rate is estimated by subtracting the effect of the discretionary policy measures, based on Istat model multipliers, from the GDP growth revised in accordance with the actual path of exogenous variables. The simulation then adds the effects of the planned discretionary fiscal measures, evaluated in terms of the multipliers for the crisis period, to come up with a GDP projection that is inclusive of the growth impact of the fiscal measures specific to the recession period. The comparison with the GDP forecast adjusted with the actual path of the exogenous variables provides an estimate of the change in the growth forecast errors attributable to the larger size of fiscal multipliers. Results are reported in Table 15.

Table 15 – Growth forecast errors based on the estimated multipliers in the crisis period
(y-o-y % changes; percentage points)

	2012	2013	2014
GDP outturn (released in March 2017) (*a*)	**-2.8**	**-1.7**	**0.1**
Istat GDP forecast (*b*) - *press-release of May 2012 (1)*	-1.5	0.6	0.6
Growth forecast errors (b - a)	*1.4*	*2.3*	*0.5*
GDP forecast (*c*) - *actual path of exogenous variables*	-1.6	-0.1	0.2
Forecast error (c - a)	*1.3*	*1.6*	*0.1*
GDP forecast (*d*) - *actual path of exogenous variables* - *baseline multipliers for the impact of fiscal shocks*	-3.0	-1.7	0.3
Forecast error (d - a)	*-0.1*	*0.0*	*0.2*
Decomposition of growth forecast errors			
Growth forecast errors (b - a)	1.4	2.3	0.5
- *forecast errors due to exogenous variables*	0.1	0.7	0.4
- *forecast errors due to fiscal multipliers*	1.4	1.6	-0.1
- *forecast errors due to other factors*	-0.1	0.0	0.2

(1) Istat economic outlook concerns the period 2012-2013. In 2014, GDP was projected to grow by 0.6 percent based on the MeMo-It model used by the PBO.

The first row of Table 15 reports the official GDP growth rate for the years from 2012 to 2014 (released on March 2017 and not subject to further revisions): Italian GDP declined by 2.8 percent in 2012 and by 1.7 percent in 2013 before increasing slightly in 2014 (0.1 percent). According to the Istat economic outlook of May 2012, Italian GDP was projected to decline by 1.5 percent in 2012 and to increase by 0.6 percent in 2013. According to our extension of the Istat macroeconomic projection by one additional year to 2014 (considering both the forecast of the exogenous variables for 2014 available at the time the Istat economic outlook was produced and the whole plan of fiscal measures), output for 2014 was projected to increase at almost the same rate as 2013 (0.6 percent).[24] The growth forecast errors are computed as the difference between the actual and the expected GDP performance (Table 15, (*b*-*a*)). This difference indicates that, in the 2012-2013 period, the recession was considerably deeper than envisaged by Istat: the fall in real output was underestimated by about 1.4 percentage points in 2012 and by 2.3 percentage points in 2013.[25] As for 2014, the prediction developed by using the MeMo-It model outpaces actual growth by 0.5 percentage points.

[24] This development is not so far from Istat GDP forecast for 2014 (first released on May 2013), which had anticipated growth of 0.7 percent.
[25] This is a common feature of the forecasting exercises performed by other institutions early in 2012. For example, consider European Commission, Spring Forecast 2012, http://ec.europa.eu/economy_finance/publications/european_economy/2012/ee1upd_en.htm; OECD Annual Projections, Economic Outlook No. 91 - June 2012, https://stats.oecd.org/Index.aspx?DataSetCode=EO91_LTB#; IMF, WEO April 2012,

As for the sources of the growth forecast errors, the simulation exercise shows that the global exogenous variables explain a minor fraction of the prediction errors, particularly in 2012 and 2013. The set of exogenous variables taken into account includes world trade, the nominal US dollar/euro exchange rate, the oil price and the nominal three-month interest rate. The most important differences between the actual path of the exogenous variables in 2012-2014 and their expected path underlying the Istat forecast regard world trade, which was overestimated by about 3 percentage points in 2013. The nominal short-term interest rate was assumed to be unchanged over the forecast horizon, whereas it actually declined to close to the zero lower bound as a result of the ECB's accommodative monetary policy. Minor discrepancies concern the US dollar/euro exchange rate, which is assumed constant over the forecast period (the observed exchange rate fell slightly in 2012 and rose in 2013), and the oil price (which was projected to decline in 2013 but at a more moderate pace compared with the actual outturn). As for 2014, we adopt assumptions for the exogenous variables in line with the forecasts available at the time: the US dollar/euro exchange rate and the short-term interest rate remain stationary; world trade growth declines by 1 percentage point compared with 2013, and the oil price declines to below 110 dollars per barrel. Overall, the expected path of the exogenous variables is more supportive of GDP growth compared with their observed development. When the actual path of the exogenous variables is taken into account, the forecast based on the Istat model is revised downwards: compared with the Istat press release, real GDP declines by 1.6 percent in 2012, becomes negative in 2013 (-0.1 percent) and increases slightly in 2014 (0.2 percent).[26] The contribution of the exogenous variables to the overall prediction error (see Table 15 (*c-a*), and the lower part of the table on the decomposition of the growth forecast error) is small in 2012 (0.1 percentage points, i.e. 7 percent of the total error) and in 2013 (0.7 percentage points, or 30 percent of the total error), but increases, considering our assumptions about the path of exogenous variables, in 2014 (0.4 percentage points, about 90 percent of the entire forecast error).

A more substantial revision of growth forecast errors is obtained when the effects of the fiscal consolidation plans are evaluated on the grounds of the multipliers specific to the crisis period (baseline multipliers). On the basis of expression (26), the fiscal measures are estimated to bring about a reduction of real GDP growth rate of 1.9 percentage points in 2012 and 2.3 percentage points in 2013, while the estimated effect is considerably lower in 2014 (0.3 percentage points).[27] The net effect of the above estimates yields a significant decrease in real GDP developments, with a GDP contraction of 3.0 percent in 2012 and 1.7 percent in 2013; the change is slightly positive

https://www.imf.org/external/pubs/ft/weo/2012/01/weodata/index.aspx.

[26] As for the Istat assumptions, the exogenous variables contributed 1 percentage point to GDP growth in 2012, 1.7 percentage points in 2013 and 1.6 percentage points in 2014. As for their actual path, the contribution is similar in 2012 (0.9 percentage points), notably lower in 2013 (1.1 percentage points) and in 2014 (1.2 percentage points).

[27] These effects compare with the much lower impact of the fiscal consolidation plan when the assessment of the same policy measures is conducted on the basis of the fiscal multipliers underlying the Istat prediction: 0.5 percentage points in 2012, 0.7 in 2013 and 0.4 in 2014.

in 2014 (Table 15, (*d*)). The growth forecast errors accounted for by the underestimation of fiscal multipliers are estimated at 1.4 percentage points in 2012 (explaining almost the entire forecast error) and 1.6 percentage points in 2013 (70 percent of the total error). For 2014, the role of fiscal multipliers is almost negligible because, as reported above, the forecast errors are primarily related to the assumptions about the exogenous variables.[28]

Overall, the former results suggest that the downturn, following the sovereign debt crisis, can be predominantly associated with the short-run effects of the fiscal consolidation plan adopted in the second half of 2011, when properly taking account of crisis-specific multipliers. This evidence is substantially in line with that of Felici et al. (2017), who show that the severe post-2011 fiscal consolidation explains a large fraction of the actual fall of GDP. It differs partly from the evidence of Busetti and Cova (2013), who estimate that the internal factors explain about two-thirds of the decline in Italian GDP but do not give strong support to the view of the larger output responses to fiscal shocks in the severe economic conditions prevailing in Italy at the time.

Finally, as a further robustness check, we assess whether the above findings might reflect difficulties in accurately measuring output multipliers. The multipliers specific to the crisis period, used to perform the above exercise (baseline multipliers), are estimated under the assumption of no correlation across subsamples of the output responses to the same fiscal shocks. The sensitivity analysis carried out in Section 5 concludes that the output responses to fiscal shocks are not significantly different from the baseline multipliers even considering several degrees of correlation. In addition, for values of ρ close to 1, the multipliers in the recession period are statistically different from the pre-crisis multipliers for the majority of fiscal instruments. In order to account for the uncertainty surrounding multiplier estimates, the simulation of the effects of the consolidation plan is replicated taking into account the estimates of fiscal multipliers for values of the correlation coefficient greater than zero. As ρ is restricted to values in the interval [0.1, 0.2,...,1], ten different sets of fiscal multipliers are used, so to obtain a range of macro-fiscal projections, GDP growth rates and growth forecast errors.

The main finding is that, as ρ increases, the magnitude of the macroeconomic effects of the fiscal measures declines compared with the estimates obtained with the baseline multipliers. The estimated impact of policy measures gives rise to a GDP fall of 1.2 percentage points in 2012 (ρ=0.1), which decreases to 0.8 percentage points (for ρ=0.9).

[28] The other factors, which are close to zero, represent the unexplained part of the decomposition. This component is mainly attributable to the following factors: *i*) the estimated full-sample multipliers could be slight different compared with the output responses to fiscal shocks implicit in the Istat forecast; *ii*) some exogenous variables are not explicitly referred to in the Istat economic outlook and are not adjusted to their actual values. This is the case of the variables pertaining to the international environment (world manufacturing export prices) as well as the indicators of uncertainty and household confidence; *iii*) it has been assumed that the fiscal measures underlying the Istat forecast were considered at their "face value", although their size could have been revised by Istat's researchers before the forecasting exercise. Furthermore, some deviations between the public finance plans considered in this study and those used by Istat are possible.

The size of fiscal effects is also revised downwards in 2013, and ranges from 1.6 (ρ=0.1) to 1.3 percentage points (ρ=0.9). The impact of policy measures also changes in 2014, although at a smaller extent (from -0.1 to 0.1). Therefore, the fraction of growth forecast errors explained by the fiscal multipliers decreases as the correlation increases. The unexplained part of the growth forecast errors ("other factors") rises to about 0.4 percentage points in 2012 (ρ=0.9), and by about the same amount in 2013, leaving further room for the effects related to factors not explicitly considered in the above simulations.

Overall, the very small values of the residual component (near zero in the case of baseline multipliers) support the assumption of larger output responses to fiscal shocks in severe economic conditions. The role of fiscal multipliers turns out to be significant in the years most affected by the effects of the consolidation measures (2012-2013), while it seems less important when the recession is milder. This is the case for 2014, when the growth forecast errors are primarily explained by the actual path of the exogenous variables.

7 Conclusions

In this paper we provide estimates of the fiscal multipliers for the Italian economy based on a structural macro-econometric model (the MeMo-It model). Performing sub-period estimations, we find suggestive evidence of an increase in the size of those multipliers in the latter part of the sample, involving the crisis period (2008-2014).

How to get from these indications to a more precise inference of crisis-specific multipliers is an unresolvable problem with standard model estimation since: *a)* model-based estimates of fiscal multipliers are, by construction, independent of the state of the economy; *b)* the length of the crisis is too short to make any accurate and efficient estimate of the crisis multipliers feasible.

We circumvent this problem by first correcting the model for any instability of structural parameters (in both cointegration relations and error correction equations) and, then, applying a methodology that, based on the estimation of weights for sub-period multipliers, allows us to infer the fiscal multipliers for the crisis period. We show that, despite the higher statistical uncertainty, the size of these multipliers is significantly larger than pre-crisis estimates, both on the expenditure and the revenue sides. This holds even after various sensitivity tests confirming, for the Italian case, the findings of the empirical literature about the notable increase in the size of fiscal multipliers in the crisis period compared with normal times.

We then apply the estimated (baseline) period-specific multipliers to the multi-year fiscal consolidation plan adopted, in several instalments, by Italy in the latter part of 2011. Our simulation results show that appropriate consideration of the crisis-specific multipliers considerably reduces the forecast error for the depth and length of the subsequent recession compared with the projection obtained on the basis of the standard model-based multipliers. Robustness checks, performed under the assumption of positive values for the covariance between the same multipliers in different periods, substantially confirm the baseline result of an appreciable reduction in the forecast error.

References

Abiad, A., Furceri, D. and Topalova, P. (2015) The Macroeconomic Effects of Public Investment: Evidence from Advanced Economies, IMF Working Paper No. 15.

Alesina A., Giavazzi F. and Favero C. (2015), “The output effect of fiscal consolidation”, *Journal of International Economics*, 96, S19-S42.

Alesina A., Giavazzi F. and Favero C. (2018), “What do we know about the effects of austerity?”, *American Economic Review Papers and Proceedings*, 108: 524-30.

Ali, M.M. and Silver, J.L. (1975), Tests for Equality Between Sets of Coefficients in Two Linear Regressions Under Heteroscedasticity, Journal of the American Statistical Association, Vol. 80, No. 391.

Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica, 59, 817–858.

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. Econometrica 61 (4), 821-56.

Arai, Y. and Kurozumi, E. (2007). Testing for the null hypothesis of cointegration with a structural break. Econometric Reviews, 26, 705–739.

Arin, KP, Koray, F. and Spagnolo, N. (2018), Fiscal multipliers in good times and bad times, Journal of Macroeconomics 44 (2015) 303–311.

Auerbach A. J. and Gorodnichenko T. (2012a), Fiscal multipliers in recession and expansion, in *Fiscal Policy after the Financial Crisis*, ed. Francesco Giavazzi and Alberto Alesina, University of Chicago Press, Chicago.

Auerbach A. J. and Gorodnichenko T. (2012b), Measuring the Output Responses to Fiscal Policy, American Economic Journal: Economic Policy 2012, 4(2): 1–27.

Bacchini et al. (2013), “Building the core of the Istat system of models for forecasting the Italian economy: MeMo-It”, Rivista di Statistica Ufficiale, Vol. 15 No. 1.

Bacchini, F., Girardi, A. and Pappalardo, C. (2015), “Macroeconometric Models in Practice: the Istat Experience”, International Journal of Computational Economics and Econometrics Vol. 6 No. 3.

Bagnai, A., Granville, B., and Mongeau Ospina, C.A., (2017). Withdrawal of Italy from the euro area: Stochastic simulations of a structural macroeconometric model, Economic Modelling, Vol. 64, 524-538.

Bai, J., and Perron, P. (1998), Estimating and Testing Linear Models With Multiple Structural Changes, Econometrica, 66, 47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models, Journal of Applied Econometrics, 18, 1–22.

Barrell R., Holland D. and Hurst I. (2012), "Fiscal multipliers and prospects for consolidation", OECD Journal: Economic Studies, Vol. 2012 No. 1.

Barro R. J. (1974), "Are Bonds net Wealth?", *Journal of Political Economy*, 82(6) 1095-1117.

Barro R. J. (1989), "The Neoclassical Approach to Fiscal Policy", in *Modern Business Cycle Theory*, ed. Robert J. Barro, 178-235, Cambridge, MA; Harvard University Press.

Batini N., Callegari G. and Melina G. (2012), "Successful austerity in the United States, Europe and Japan", IMF working paper 12/190.

Batini N., Eyraud L., Weber A. (2014a), "A Simple Method to Compute Fiscal Multipliers" IMF Working paper 14/93.

Batini N., Eyraud L., Forni L. and Weber A, (2014b), "Fiscal multipliers: size, determinants and use in macroeconomic projections", IMF Technical Notes and Manuals, September.

Baum A. and Koester, G.B. (2011), The impact of fiscal policy on economic activity over the business cycle - evidence from a threshold VAR analysis, Deutsche Bundesbank Discussion Paper Series 1: Economic Studies No 03.

Baum A., Poplawski-Ribeiro M., Wener A. (2012), "Fiscal multipliers and the state of the economy", IMF working paper 12/286.

Beyer, A., Dewald, William G. and Haug, Alfred A., (2009). Structural breaks, cointegration and the Fisher effect, Working Paper Series No. 1013, European Central Bank.

Blanchard O., Dell'Ariccia G. and Mauro P. (2010) "Rethinking Macroeconomic Policy", Journal of Money, Credit and Banking 42(s1), 199-215.

Blanchard O. and Leigh D. (2013), "Growth Forecast Errors and Fiscal Multipliers", *American Economic Review: Papers and Proceedings* 103(3) 117-120.

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.

Boitani, A. and Perdichizzi, S. (2018), “Public Expenditure Multipliers in Recessions. Evidence from the Eurozone”, Working Paper Series n. 68, Università Cattolica del Sacro Cuore.

Boldea, O., Hall A. R., and Han S. (2012). “Asymptotic distribution theory for break point estimators in models estimated via 2SLS”, Econometric Reviews 31: 1-33.

Borenstein, M., Hedges, L., Higgins, J. and Rothstein, R. (2009), Introduction to Meta-Analysis, Wiley.

Brown, R.L., Durbin, J. And Evans, J.M. (1975), Techniques for testing the constancy of regression relationships over time. Journal of the Royal Statistical Society B 37, 149-163.

Busetti, F. and Cova, P. (2013), L’impatto macroeconomico della crisi del debito sovrano: un’analisi controfattuale per l’economia italiana, Questioni di economia e finanza No. 201, Bank of Italy.

Caggiano G., Castelnuovo E., Colombo V. and Nodari G. (2015), “Estimating Fiscal Multipliers: News From A Non-linear World”, Econ J, Vol. 125 No. 584.

Caprioli, F. and Momigliano, S. (2011). The effects of fiscal shocks with debt-stabilizing budgetary policies in Italy, Temi di discussione (Economic working papers) No. 839, Bank of Italy.

Caprioli, F. and Momigliano, S. (2013). The Macroeconomic Effects of Expenditure Shocks During Good and Bad Times, in *Fiscal Policy and Growth*, Workshop and Conferences, Bank of Italy, June 2013.

Carreras O., Kirby S., Liadze I. and Piggott R. (2016) “Quantifying Fiscal Multipliers”, NIESR Discussion Paper No. 469.

Carrion-i-Silvestre, J. L., and Sansó, A. (2006), Testing the Null Hypothesis of Cointegration With Structural Breaks, Oxford Bulletin of Economics and Statistics, 68, 623–646.

Casella G. and Berger R. L. (2002), Statistical inference, Duxbury Thomson Learning.

Chinn M. D. (2013), “Fiscal multipliers” the *New Palgrave Dictionary of Economics.*

Chow, G. C. (1960), Tests of Equality Between Sets of Coefficients in Two Linear Regressions, Econometrica, Vol. 28, No. 3.

Christiano L., Eichenbaum M. and Rebelo S. (2011), “When Is the Government Spending Multiplier Large?” *Journal of Political Economy* 119(1) 78-121.

Cimadomo J. and D'Agostino A. (2015), “Combining time-variation and mixed-frequencies: an analysis of government spending multipliers in Italy”, Journal of Applied Econometrics, Vol. 31.

Coenen G., Erceg C., Freedman C., Furcei D., Kumhof M., Lalonde R.; Laxton D., Lindé J., Mourougane A., Muir D., Mursula S., de Resende C., Roberts J., Roeger W., Snudden S., Trabandt M., in 't Veld J. (2012), “Effects of Fiscal Stimulus in Structural Models”, *American Economic Journal: Macroeconomics* 4(1): 22–68.

Cogan J. F., Cwick T., Taylor J. B. and Wieland V. (2010), “New Keynesian versus Old Keynesian Government Spending Multipliers”, *Journal of Economic Dynamics and Control*, 34 281-295-

Cole H. L. and Ohanian L. E. (2004), “New Deal Policies and the Persistence of the Great Depression; A General Equilibrium Analysis”, *Journal of Political Economy* 112(4) 779-816.

Corsetti G., Meier A. and Muller G. (2012), “What determines government spending multipliers”, Economic Policy, Vol. 27, No. 72.

Dolls M., Fuest C. and Peichl A. (2012), “Automatic Stabilizers and Economic Crisis: Us vs Europe”, *Journal of Public Economics* 96 279-294.

Draper, D. et al. (1993), On Combining information. Statistical issues and opportunities for research, US National Research Council’s Committee on Applied and Theoretical Statistics.

Dülger, F. (2016). The Sustainability of Current Account in the Presence of Endogenous Multiple Structural Breaks: Evidence from Developed and Developing Countries, Panoeconomicus, Vol. 63/3.

Eggertsson, G. B. **(**2009) What Fiscal Policy Is Effective at Zero Interest Rates?, Federal Reserve Bank of New York Staff Report 402.

Fair R. C. (1984), Specification, Estimation and Analysis of Macroeconometric Models, Harvard University Press.

Fair R. C. (2004), Estimating How the Macroeconomy Works, Harvard University Press.

Fair R. C. (2015), “Reflections on macroeconometric modeling”, The B.E. Journal of Macroeconomics, Vo. 15 N. 1.

Fatas A., Mihov I. (2012),”Fiscal Policy as a Stabilization Tool” The B.E. Journal of Macroreconomics, Special issue: Long-Term Effects of the Grat Recession, 12-3, 1935-1690.

Favero, C. and Giavazzi, F. (2009), Measuring Tax Multipliers: The Narrative Method in Fiscal VARs, American Economic Journal: Economic Policy 2012, 4(2).

Fazzari S. M., Morely J. and Panovska I. (2015), State dependent effects of fiscal policy, Studies in Nonlinear Dynamics and Econometrics, Vol. 19, No. 3.

Felici F., Nucci F., Ricchi O. and Tegami C. (2017), "The Post-2007 Developments in the Italian Economy: A Counterfactual Analysis with the ITEM Model", in Paganetto L. (ed.) Sustainable Growth in the EU. Challenges and Solutions, Springer.

Fisher, F.M. (1970), Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An ExpositoryNote, Econometrica, Vol. 38, No. 2

Figueres J. M. (2015), "News on state dependent fiscal multipliers: the role of confidence", Università degli Studi di Padova, Marco Fanno working paper n. 202. September.

Gardner, L.A. (1969), On detecting changes in the mean of normal variates. The Annals of Mathematical Statistics 40, 116-126.

Gechert S., Huges Hallet A., Rannemberg A. (2016), "Fiscal multipliers in downturns and the effects of Euro Area consolidation", Applied Economics Letters 23(16) 1138-1140.

Gechert, S. (2015) " What fiscal policy is most effective? A meta-regression analysis", Oxford Economic Papers, 67(3), 2015, 553–580.

Ghysels, E., Guay, A. and Hall, A. (1998). Predictive Tests for Structural Change with Unknown Breakpoint. Journal of Econometrics 82 (2), 209-233.

Górnicka, L., Kamps, C., Köster, G., and Leiner-Killinger, N. (2018), " Learning about fiscal multipliers during the European sovereign debt crisis: evidence from a quasi-natural experiment", ECB Working Paper No. 2154.

Gregory, A.W. and Hansen, B.E. (1996). Residual-based tests for cointegration in models with regime shifts. Journal of Econometrics 70, 99-126.

Guajardo J., Leigh D. and Pescatori A. (2014), "Expansionary Austerity? International evidence", Journal of European Economic Association, 12(4) 948-968.

Gupta, S. A. (1978), "Testing the Equality Between Sets of Coefficients in Two Linear Regressions When Disturbances Are Unequal," unpublished Ph.D. dissertation, Purdue University, Dept. of Statistics.

Gupta, S.A. (1982), "Structural Shift: A Comparative Study of Alternative Tests," in Proceedings of the Business and Economic Statstics Section, American Statstical Association, pp. 328-331.

Hall, A. R., Han, S. and Boldea, O. (2012). Inference regarding multiple structural changes in linear models with endogenous regressors, Journal of Econometrics, Vol. 170/2: 281-302.

Hall, A. R., Osborn, D. R., and Sakkas, N. (2015). Structural Break Inference Using Information Criteria in Models Estimated by Two-Stage Least Squares. Journal of Time Series Analysis,36: 741–762.

Hatemi-J, A., (2008). Tests for cointegration with two unknown regime shifts with an application to financial market integration. Empirical Economics 35, 497–505.

Haug, A.A., and Siklos, P. (2006), The Behavior of short-term interest rates: International evidence of nonlinear adjustment, Studies in Nonlinear Dynamics and Econometrics, 10, Article 6, 1–32.

Hedges, L., and Olkin, I. (1985). Statistical Methods for Meta-analysis. San Diego, CA: Academic Press.

Hendry, D.F. (1999). An econometric analysis of US food expenditure, 1931-1989. In Magnus, J.R. and Morgan, M.S. (eds) Methodology and tacit knowledge: Two experiments in econometrics. p.341-361. New York: Wiley.

Hendry D. F., A. R. Pagan and Sargan J. D. (1984), “Dynamic Specification”, in Z. Griliches e M. D. Intriligator (eds.), Handbook of Econometrics, Vol. II, North Holland.

Hendry, D.F. Johansen, S. and Santos, C. (2008). Automatic Selection of Indicators in a Fully Saturated Regression. Computational Statistics 23, 317–335.

Hsiao C. (1997), “Cointegration and Dynamic Simultaneous Equations Model”, Econometrica, Vol. 65 N. 3.

Hunter, J.E., and Schmidt, F.L. (2004). Methods of Meta-analysis: Correcting Error and Bias in Research Findings (2nd edn). Newbury Park, CA: Sage Publications.

Ilzetzki E., Reinhart C.M., and Rogoff K.S. (2009), Exchange Rate Arrangements Entering the 21st Century: Which Anchor Will Hold? Mimeo. University of Maryland and Harvard University.

Iltzetzki E., Mendoza E. G. and C. A. Végh (2013), “How big (small?) are fiscal multipliers?”, Journal of Monetary Economics 60(2) 239-254.

Jayatissa, W. A. (1977), Tests of Equality between Sets of Coefficients in Two Linear Regressions When Disturbance Variances are Unequal. Econometrica 45: 1291-1292.

James, W. and Stein, C. (1961), Estimation with quadratic loss, Proc. Fourth Berkeley Symposium, 1, 361–380, Univ. California Press.

Johansen, S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press.

Jorda O. and Taylor A. M. (2016), "The time for austerity: estimating the average treatment effect of fiscal policy", The Economic Journal, Vol. 126, No. 590.

Karras, G. (2014), Is Fiscal Policy More Effective During Cyclical Downturns?, International Economic Journal, 28:2.

Kejriwal, M. (2008), Cointegration with structural breaks: An application to the Feldstein–Horioka Puzzle. Studies in Nonlinear Dynamics & Econometrics, 12, 1–37.

Kejriwal, M., and Perron, P. (2008a), Data Dependent Rules for the Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression, Econometric Theory, 24, 1425–1441.

Kejiriwal, M., and Perron, P. (2008b), The limit distribution of the estimates in cointegrated regression models with multiple structural changes, Journal of Econometrics, 146, 59–73.

Kejiriwal, M., and Perron, P. (2010), Testing for multiple structural changes in cointegrated regression models, Journal of Business and Economic Statistics, 28, 503–522.

Keller, T. and Holkin, I. (2004), Combining correlated unbiased estimator of the mean of a normal distribution, A Festschrift for Herman Rubin, Institute of Mathematical Statistics, Lecture notes, Monograph series, Vol. 45, Beachwood Ohio.

Kilponen, J. et al. (2015), Comparing fiscal multipliers across models and countries in Europe, ECB Working Paper, No. 1760.

Kirchner M., Cimadomo J. and Hauptmeier S. (2010), "Transmission of Government Spending Shocks in The Euro Area: Time Variation and Driving Forces", ECB working papers series 1219.

Kubokawa, T. (1999), Shrinkage and modification techniques in estimation of variance and the related problems: A review, Communications in Statistics - Theory and Methods.

Locarno, A. Notarpietro, A. and Pisani, M. (2013), Sovereign risk, monetary policy and fiscal multipliers: a structural model-based assessment, Temi di discussione, No. 943, Bank of Italy.

Lütkepohl H. (2005), New Introduction to Multiple Time Series Analysis, Springer Verlag.

Liu. J, Wu, S. and Zidek, J.V. (1997), On Segmented Multivariate Regressions, Statistica Sinica, 7, 497-525.

Maatta, J.M. and Casella, G. (1990), Developments in Decision-Theoretic Variance Estimation, Statistical Science, 5, 90-101.

Maki, D. (2012). Tests for cointegration allowing for an unknown number of breaks, Economic Modelling, Vol. 29, 2011-2015.

Mertens, K. and Ravn, M. O. (2013), The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, American Economic Review, 103 (4): 1212-47.

Mineshima A., Poplawski-Ribiero M. and Weber A. (2014), "Size of Fiscal Multipliers", in Post-Crisis Fiscal Policy, ed. Carlo Cottarelli, Philip Gerson and Abdelhak Senhadji, The MIT Press Cambridge Massachussets.

Mountford A. and Uhlig H. (2009), "What are the effects of fiscal policy shocks?", Journal of Applied Econometrics, 24(6) 960-992.

Nakamura E. and Steinsson S. (2014), "Fiscal stimulus in a monetary union: Evidence from U.S. regions", *American Economic Review*, 104 (3): 753-92..

Ohtani, K. and Kobayashi, M. (1979), A Bounds Test for Equality between Sets of Coefficients in Two Linear Regression Models under Heteroscedasticity, Econometric Theory, Vol. 2, No. 2.

PBO, Budgetary Planning Report, 2017.

Perotti R. (2013), "The Austerity Mith: Gain without Pain", in Fiscal Policy after the Great Recession, ed. Alberto Alesina and Francesco Giavazzi, University of Chicago Press, Chicago 307-354.

Perron, P. (2006), Dealing With Structural Breaks, in Palgrave Handbook of Econometrics, eds. K. Patterson and T. C. Mills, New York: Palgrave Macmillan, pp. 278–352.

Perron, P. and Yamamoto, Y. (2015), Using OLS to Estimate and Test for Structural Changes in Models with Endogenous Regressors, Journal of Applied Econometrics ,Vol. 30, 119–144.

Phillips, P. C. B. and Ouliaris, S. (1990), Asymptotic properties of residual based tests for cointegration. Econometrica, Vol. 58, pagg. 165–193.

Quandt, R.E., (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. Journal of the American Statistical Association 55, 324-330.

Ramey V. A. and Zubairi S. (2018), "Government Spending Multipliers in Good times and in Bad: Evidence from US Histrorical Data", Journal of Political Economy, Vol. 126, No. 2.

Rannenberg, A., Schoder, C. and Strasky, J. (2015), The macroeconomic effects of the Euro Area's fiscal consolidation, Economic Letters 10/EL/15, Central Bank of Ireland.

Romer, C. D., and Romer, D. H. (2010), The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks, American Economic Review, 100, 763–801.

Sanchez-Meca, J., and Marın-Martınez, F. (1998), Weighting by inverse variance or by sample size in meta-analysis: A simulation study. Educational and Psychological Measurement, 58, 211-220.

Sanchez-Meca, J. and Marin-Martinez, F. (2010), Weighting by Inverse Variance or by Sample Size in Random-Effects Meta-Analysis, Educational and Psychological Measurement, 70(1).

Schmidt P. (1973), The asymptotic distribution of dynamic multipliers, Econometrica, Vol. 41.

Schmidt, F. L., Oh, I.-S., and Hayes, T. L. (2009), Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of difference in results. British Journal of Mathematical and Statistical Psychology, 62, 97-128.

Schmidt, P. and Sickles, R. (1977) Some Further Evidence on the Use of the Chow Test under Heteroscedasticity. Econometrica 45, 1293-1298.

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. Econometric Theory, 10, 91–115.

Spilimbergo A., Symarsky S. and Schindler M., “Fiscal multipliers”, IMF Staff Position Note, 09/11, May.

Theil, H. (1971), Principles of Econometrics, New York: John Wiley.

Toyoda, T. (1974), Use of the Chow Test under Heteroscedasticity. Econometrica. 42: 601-608.

Tong, J. and Wang, Y. (2007), Optimal Shrinkage Estimation of Variances with Applications to Microarray Data Analysis, JASA.

Tong, J. and Wang, Y. (2012), James Stein type estimators of variances, Journal of Multivariate Analysis.

Warmedinger, T. Checherita-Westphal, C. and Hernández de Cos, P. (2015), Fiscal multipliers and beyond, ECB Occasional Paper, No. 162.

Watt, P. A. (1979), Tests of Equality between Sets of Coefficients in Two Linear Regressions When Disturbance Variances are Unequal: Some Small Sample Properties. The Manchester School 47: 391-396.

Woodford M. (2011), Simple Analytics of the Government Expenditure Multiplier, American Economic Journal: Macroeconomics 3(1) 1-35.

Zellner, A. (1962), An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, Journal of the American Statistical Association, 57, 348-368.