IMK
Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

Juli 2012
97

Working Paper

Sebastian Gechert* | Henner Will**

Fiscal Multipliers: A Meta Regression Analysis

Abstract
Since the fiscal expansion during the Great Recession 2008-2009 and the current European consolidation and austerity measures, the analysisof fiscal multiplier effects is back on the scientific agenda. The number of empirical studies is growing fast, tackling the issue with manifold model classes, identification strategies, and specifications. While plurality of methods seems to be a good idea to address a complicated issue, the results are far off consensus. We apply meta regression analysis to a set of 89 studies on multiplier effects in order to provide a systematic overview of the different approaches, to derive stylized facts and to separate structural from method-specific effects. We classify studies with respect to type of fiscal impulse, model class, multiplier calculation method and further control variables. Moreover, we analyse subsamples of the model classes in order to evaluate the effects of model-class-specific properties, currently discussed in the literature, such as the influence of central bank reaction functions and liquidity constrained households. As a major result, we find that the reported size of the fiscal multiplier crucially depends on the setting and method chosen. Thus, economic policy consulting based on a certain multiplier study should lay open by how much specification affects the results. Our meta analysis may provide guidance concerning influential factors.

Keywords: multiplier effects; fiscal policy; meta analysis.

JEL ref.: E27, E62, H30.

---

* Chemnitz University of Technology. Corresponding author, gechert@wirtschaft.tu-chemnitz.de
** Macroeconomic Policy Institute. My good friend and coauthor Henner Will passed away in a tragic road accident on November 18, 2011 at the age of 28.

Fiscal Multipliers: A Meta Regression Analysis*

Sebastian Gechert[†] Henner Will[‡]

July 2012

Abstract. Since the fiscal expansion during the Great Recession 2008-2009 and the current European consolidation and austerity measures, the analysis of fiscal multiplier effects is back on the scientific agenda. The number of empirical studies is growing fast, tackling the issue with manifold model classes, identification strategies, and specifications. While plurality of methods seems to be a good idea to address a complicated issue, the results are far off consensus. We apply meta regression analysis to a set of 89 studies on multiplier effects in order to provide a systematic overview of the different approaches, to derive stylized facts and to separate structural from method-specific effects. We classify studies with respect to type of fiscal impulse, model class, multiplier calculation method and further control variables. Moreover, we analyse subsamples of the model classes in order to evaluate the effects of model-class-specific properties, currently discussed in the literature, such as the influence of central bank reaction functions and liquidity constrained households. As a major result, we find that the reported size of the fiscal multiplier crucially depends on the setting and method chosen. Thus, economic policy consulting based on a certain multiplier study should lay open by how much specification affects the results. Our meta analysis may provide guidance concerning influential factors.

Keywords. multiplier effects; fiscal policy; meta analysis

JEL classification. E27, E62, H30

*We are grateful to Johanna Kemper, Rafael Wildauer and Michael Hachula for excellent research assistance. We would like to thank the authors of the studies we analysed for providing us with additional information. Moreover, we are indebted to Fritz Helmedag, the whole Macroeconomic Policy Institute research team, the participants of the Euroframe conference in Kiel 2012, INFER annual conference in Coimbra 2012, Keynes Society meeting in Linz 2012 and Nifip conference in Porto 2011 for their helpful comments. Sebastian Gechert would like to gratefully acknowledge financial support by Macroeconomic Policy Institute.

[†]Chemnitz University of Technology. Corresponding author, gechert@wirtschaft.tu-chemnitz.de

[‡]Macroeconomic Policy Institute (IMK). My good friend and coauthor Henner Will passed away in a tragic road accident on November 18, 2011 at the age of 28.

1 Introduction

The discussion on the scale of fiscal multipliers has lasted for decades and still economists struggle on the value of the multiplier. Stimulus packages facing the Great Recession and the current consolidation and austerity measures have brought the matter back on the scientific agenda. Especially, the question of effects of the US stimulus packages under Bush jun. and Obama administrations were a permanent source of economic discussion. Turning to European countries, the effects of fiscal contractions are a central issue that is closely related to multiplier evaluations. In theoretical approaches several effects have been discussed that eventually turn the balance of the multiplier below or above unity. Roughly summing up, the discussion is about crowding in vs. crowding out effects in private consumption, investment and net exports.

The empirical literature on the size of the multiplier is growing fast, tackling the issue with manifold model classes, identification strategies, and specifications. While plurality of methods seems to be a good idea to address a complicated issue, unsurprisingly the results are far off consensus.

The vast majority of different model approaches and assumptions makes the case for a systematic literature review. Several papers that try to summarize the literature take a descriptive approach or come up with a list of reported multipliers and characteristics of the reporting studies. However, since reported multiplier values in the literature are quantifiable, it should be possible to review the literature with statistical criteria. Meta regression analysis is a suitable tool to tackle the issue. According to Stanley and Jarrell (2005: 301), 'Meta analysis is the analysis of empirical analyses that attempts to integrate and explain the literature about some specific important parameter.'

We apply meta regression analysis to a set of 89 studies on multiplier effects in order to provide a systematic overview of the different approaches, to derive stylized facts and to separate structural from method-specific effects.

It should be stressed that our method is not suitable to find *the true multiplier value*, because even if our sample is an unbiased representation of the whole literature on multiplier effects, it is not clear whether or not this whole literature provides an unbiased picture of actual multiplier effects. Moreover, as Carroll (2009: 246) points out, 'asking what *the* government spending multiplier is, [...] is like asking what *the* temperature is. Both vary over time and space.' However, our meta analysis helps to filter out the systematic influence of certain study characteristics on the reported multiplier value. We are able to separate methodic distinctions among studies from structural distinctions of the fiscal policy settings these studies evaluate.

We classify studies with respect to type of fiscal impulse, model class, multiplier calculation method and some further control variables. The type of fiscal impulse is our central structural characteristic by which we try to identify the relative effectiveness of different fiscal measures. The model class is our central method-specific parameter by which we try to analyse the goodness of fit of results from certain model classes in comparison to other model classes. Moreover, we analyse subsamples of the model classes in order to evaluate the effects of model-class-specific properties currently discussed in the literature, such as the influence of central bank reaction functions, liquidity constrained households or the sample period on which the studies base their calculations.

Our main results go in line with theoretical reasoning: first, reported multipliers largely depend on model classes, with RBC models standing out of the rest of approaches by reporting significantly lower multipliers. Second, direct public demand tends to have higher multipliers than tax cuts and transfers. Especially public investment seems to be the most effective fiscal impulse. Third, reported multipliers strongly depend on the method and horizon of calculating them. Thus, a simple listing of multiplier values without additional information on how they were computed could show a biased picture. Fourth, longer time series tend to imply higher multipliers in our sample, time series that end in more recent years tend to imply lower multipliers. One should, however, be aware that even the most recent time series in our sample do not cover an adequate portion of the effects of the stimulus packages in response to the Great Recession. Fifth, the more open the import channel of an economy, the lower seems to be the multiplier. Sixth, in model based approaches the interest rate reaction function is a key parameter to the reported multiplier value. Multiplier effects are highest, when the central bank accommodates fiscal policy or is bound to a zero interest rate. Seventh, an increasing share of Keynesian agents, for whom Ricardian equivalence is broken, significantly increases multiplier values. Both an accommodating monetary policy and liquidity constrained households correspond to the current macroeconomic setting which could imply a higher effectiveness of fiscal policy in times of the current crisis. Eighth, divergent results of the various identification strategies in VAR models seem to be partly due to multiplier calculation methods and horizons of measurement that reflect different shapes of impulse response functions.

To sum up, reported multipliers very much depend on the setting and method chosen, thus, economic policy consulting based on a certain multiplier study should lay open by how much specification affects the results. Our meta analysis may provide guidance concerning such influential specifications and their direction of influence.

The paper is organized as follows: In the next section we provide a conventional literature review on related multiplier surveys, meta analyses as well as on the topics discussed in the fiscal multiplier literature. Section three gives an overview of the data collection and variables. Section four shows some descriptive statistics. Section five explains and discusses the meta regression method. Section six provides the findings of our meta regression, including various robustness checks. The final section concludes.

2 Literature Review

2.1 Other Meta Analyses and Multiplier Literature Reviews

The growing interest in the effects of fiscal policy measures has recently provoked several overview articles that descriptively sum up the findings in the literature by extracting some stylized facts and influences of the economic setting and study characteristics (Ramey 2011; Parker 2011; Hebous 2011; Bouthevillain et al. 2009; Spilimbergo et al. 2009; van Brusselen 2009; Fatás and Mihov 2009; Hasset 2009). While at least some of these studies provide tables of study results and study characteristics to categorize the existing literature, there is a lack of a systematic quantitative analysis, which makes the case for a meta regression analysis.

Meta analysis is becoming a more and more accepted tool in economics. Ebscohost shows more than 250 entries with the phrase "meta analysis" in the title of the publication by the end of 2011. To our knowledge, our study is the first application of meta regression analysis to the growing literature on fiscal multipliers. There are some similar studies on other macroeconomic policy evaluations. De Grauwe and Costa Storti (2004) meta analyse the effects of monetary policy on growth and prices. They draw on 43 empirical studies that use VAR models and structural econometric models. Rusnák et al. (2011) reveal study specific influences, sufficient to explain the price puzzle in a sample of 70 papers on price effects of monetary policy. Another meta analysis by Nijkamp and Poot (2004) surveys 93 studies on fiscal policy, but focuses on long-run growth effects of fiscal policies, and does not take into account short-run multiplier effects. Card et al. (2010) analyse 97 studies on active labor market policies and evaluate the effectiveness of certain kinds of programs. The famous meta analysis of Card and Krueger (1995) provided insights of the reported effects of minimum wages depending on the study specification. Feld and Heckemeyer (2011) meta analyse 45 studies on the tax semi-elasticity of FDI. An overview on some further meta studies in economics can be found in Stanley (2001: 134).

2.2 Overview of Fiscal Multiplier Literature

When looking at fiscal multiplier effects, the paramount distinction concerns the types of fiscal impulses that the studies evaluate. We identified six expansionary fiscal measures, namely public consumption, public investment, military spending, direct public employment, transfers to households and tax cuts, notwithstanding more detailed classifications. Many studies do not even distinguish public consumption, investment and military spending, but simply refer to public spending without any specification.

Results are very mixed and it is not easy to find prima facie evidence concerning the relative effectiveness of fiscal measures. While some studies prefer direct public spending, some find higher multipliers for tax reliefs or transfers, and others point to a high impact of military expenses on growth. Our hypothesis is that the wide variety is largely due to a lot of interfering study characteristics, whose particular impact we try to identify via meta regression analysis.

The most prominent study characteristic is the model class. Our survey includes model-based studies as well as empirical investigations. We discriminate between New Classical RBC (or D(S)GE) models, New Keynesian DSGE models, structural macroeconometric models, VAR models, and all kinds of single equation estimation techniques (OLS, IV, ML, GMM, ECM, ...).

Basic RBC models entail a utility maximizing, representative household for whom Ricardian equivalence holds. Additionally, they feature fully competitive labor and goods markets. These models imply full crowding out of private consumption. Expansionary fiscal policy does not increase GDP via a Keynesian demand effect, but via a neoclassical negative wealth effect that results in increased labor supply (Baxter and King 1993). The multiplier effect of public spending is usually in a range of $0 < k < 1$, with the precise value depending on the elasticities of demand for labor and the elasticity of substitution of consumption and leisure (Woodford 2011). Some modifications to the household's utility function, such as complementarity of consumption and labor supply, complementarity of public and private consumption or allowing for productivity enhancing effects of public spending, may raise the multiplier to values larger than one (Linnemann 2006; Mazraani 2010). Negative multipliers in these models may come with public employment lowering private labour supply and with distortional effects of taxation (Ardagna 2001; Fatás and Mihov 2001).

Most contemporary studies on fiscal multipliers use New Keynesian DSGE models (henceforth: DSGE-NK), extending the standard RBC model with monopolistic competition producing sticky prices and/or wages. These New Keynesian amendments allow for an output gap in the short run and possible demand side effects of fiscal policy, even if Ricardian equivalence holds. Multiplier effects in these models, however, largely depend on the reaction function of the monetary authority, or more precisely on the reaction of the real interest rate. The usual setting of an inflation target or some sort of Taylor rule implies a counteraction to a decreasing output gap leading to a partial interest rate crowding out of investment and/or consumption. Depending on calibration and/or estimation of the parameters, the multiplier effects in these models vary slightly, but they typically find multipliers of public spending in a range of $0 < k < 1$. However, current developments in the related literature tend to broaden the spectrum of possible multipliers in both directions. On the one hand, the multiplier may be $k < 0$ when including so-called non-Keynesian effects due to distortionary taxation, a wage-level increasing effect of public employment, or risk premia on interest rates for high government debt. The modifications possibly indicate expansionary effects of fiscal contractions in these models (Briotti 2005: 10-11). On the other hand, introducing a share of non-Ricardian consumers (Galí et al. 2007; Cwik and Wieland 2011), or a central bank that operates at the zero lower bound (ZLB) (Woodford 2011; Freedman et al. 2010), DSGE-NK models yield higher multiplier values, comparable to those of structural macroeconometric models. Ricardian equivalence is broken by assuming high individual discount rates or liquidity constraints for some households. There are many synonyms in the literature, e. g. non-Ricardian agents, hand-to-mouth consumers, myopic agents, rule-of-thumbers, liquidity constrained households, etc. They are subsumed under the heading of *Keynesian agents* here, as they share the attribute of aligning their spending with current income. The ZLB effect constitutes a non-linearity to the central bank reaction function in situations with a big output gap and low inflation. At the ZLB the nominal interest rate is fixed, and thus expansionary fiscal policy lowers the expected real rate of interest due to increasing inflation expectations, i. e. a Fisher effect.

A third type of models are structural macroeconometric models (henceforth: MACRO), still in use for political consulting despite the dominance of micro-founded models in academia. Macroeconometric models typically do not incorporate utility maximising households, but estimate macroeconomic consumption and investment functions. Most of these models combine Keynesian reactions in the short-run with neoclassical features in the long run. Due to the short-term nature of fiscal multiplier measures their Keynesian features are core here, which usually leads to multipliers larger than one due to crowding in of private consumption or investment, depending on the monetary and foreign trade regime.

Another strand of the literature applies VAR models and measures impulse-responses of fiscal shocks. Estimated multiplier values vary widely, which may be due to divergent data bases, kinds of fiscal shocks, and the method of identification of exogenous fiscal shocks. There are five established approaches for identification, two of which rely on additional historical information, and three of which try to identify exogenous fiscal shocks directly from the time series (see Caldara and Kamps (2008) for a comprehensive explanation of most of the methods). (1) The war episodes approach focuses on a few periods of extraordinary US military spending hikes, which are deemed to be orthogonal to business cycle fluctuations (Ramey and Shapiro 1998). (2) The so-called narrative record, established by Romer and Romer (2010), follows a similar idea, but employs real time information such as government announcements or economic forecasts, and is not limited to military spending. (3) The recursive VAR approach (Fatás and Mihov 2001) uses a Choleski decomposition with imposed zero restrictions to implement a causal order of the VAR variables and to rule out contemporaneous reactions of the fiscal variable to business cycle variations. (4) The Blanchard and Perotti (2002) SVAR approach builds on the recursive VAR approach, but additionally allows for non-zero restrictions such as imposing estimated elasticities of automatic stabilizers.(5) The sign restricted VAR approach (Mountford and Uhlig 2009) identifies exogenous fiscal shocks by imposing sign restrictions to the impulse-response functions of the fiscal shocks and then distinguishing them from a business cycle shock. Some VAR studies additionally distinguish multiple regimes in order to separate effects of fiscal policy in upturns and downturns, pointing out the relevance of downturn regimes when it comes to evaluating fiscal stimuli (Auerbach and Gorodnichenko 2012).

Our data set also includes a group of various single equation estimations (henceforth: SEE), such as OLS, IV, ML, GMM, and ECM approaches. Just like VAR studies, this group reports a wider range of multiplier values than the more model based approaches. The multiplier in single equation estimations usually appears in the coefficients of the (lagged) fiscal variables, which may impose a problem to compare these multipliers with those of the other approaches.

Besides model classes, another method based characteristic is simply the way of calculating the multiplier value. In general, multiplier values are drawn from standardized fiscal impulses (e. g. 1 percent of GDP or 1 currency unit) that allow for a dimensionless comparable input-output relation. It is basically assumed that multiplier effects are independent of scale. As opposed to the comparative static textbook multiplier, dynamic and empirical approaches allow for several variants and require additional information to calculate the effect size. In line with Spilimbergo et al. (2009: 2) we found several calculation methods of the multiplier in the literature. DSGE, RBC, macroeconometric and VAR models usually provide impulse response functions of standardized fiscal policy shocks. Multipliers are calculated either as the *peak* response of GDP with respect to the initial fiscal impulse

$$k = \frac{\max_n \Delta Y_{t+n}}{\Delta G_t} \tag{1}$$

or as the *integral* of the response function of GDP divided by the integral of the fiscal impulse function

$$k = \frac{\sum_n \Delta Y_{t+n}}{\sum_n \Delta G_{t+n}} \tag{2}$$

or as the *impact* response divided by the impact impulse

$$k = \frac{\Delta Y_t}{\Delta G_t}. \tag{3}$$

Equations show that an additional information concerning the horizon of measurement n is needed. The calculation method and the parameter n may have a big impact on multipliers, especially when impulse and response functions are not smooth. Since peaks are usually the maxima of response functions, we would expect peak multipliers to exceed integral multipliers. However, sharply declining fiscal impulse functions combined with long-lasting GDP responses can produce integral multipliers exceeding peak multipliers. Impact multipliers can be subsumed under integral multipliers with a horizon of $n = 0$. For single equation estimations multiplier effects show up in the coefficients of the exogenous fiscal variables, as long as the dependent variable and the fiscal variable have the same dimension. Horizons can also be recorded via the number of lagged fiscal variables.

3 Data and Variables

Our data set includes empirical and (semi-)calibrated papers on short-term output effects of discretionary fiscal policy measures. It takes into account 89 papers from 1992 to 2012, providing a sample of 749 observations of multiplier values. We counted 278 observations from DSGE-NK models, 55 from RBC models, 94 from MACRO models, 260 from VARs and 62 from SEE. The majority of papers in our sample has been published from 2007 onwards. This is due to the fact that fiscal policy is back on the political agenda since the Great Recession.

In order to search for papers we used BusinessSearch and repec as well as established working paper series (NBER, CEPR, IMF, Fed, ECB) and Google Scholar. As a necessary precondition papers must provide calculations of multiplier effects or at least provide enough information such that we were able to calculate multiplier effects on our own. For example, some papers provided elasticities of output with respect to government spending. If these papers also provided the share of government spending to GDP, multiplier calculations were possible.

The 749 reported multiplier values come along with specific characteristics. We developed a set of characteristics that should explain the variability in the reported multiplier values. To this end, we focussed on typical characteristics that gave rise to discussions in the literature. However, some characteristics do not apply to every model class. For example, it is not possible to discriminate agent behavior in VAR studies. Thus, for the total sample we only included characteristics that fit to all model classes. In subsamples that focus on special model classes we were able to check the influence of further characteristics.

Most characteristics, such as the model class itself, are measurable on a nominal scale only, i. e. there is no possible ranking order. We group these characteristics, since they are exclusive. A reported multiplier value must exclusively belong to one value in the group 'model class', which comprehends the values (RBC, DSGE-NK, MACRO, VAR, SEE). For example, an observation that stems from a VAR has dummies (RBC=0, DSGE-NK=0, MACRO=0, VAR=1, SEE=0).

For the total sample we focus on the influence of model classes and the type of fiscal impulse (SPEND, CONS, INVEST, MILIT, TRANS, EMPLOY, TAX), which is recorded on a nominal scale, too. Again, an observation must belong to exclusively one value in this group. The value 'SPEND' applies, when the paper reports the effect of public spending without specifying whether it is public consumption (CONS), public investment (INVEST) or military spending (MILIT). Other impulses could be transfers to households (TRANS), public employment (EMPLOY) or lowering taxation (TAX).[1] For robustness checks, we also set up a variable for spending in general (GSPEND), comprising all observations from (SPEND, CONS, INVEST, MILIT), as opposed to the other types of impulses.

Moreover, we include some control variables. We record the group of multiplier calculation methods with the variables (PEAK, INTEGRAL, COEFF) on a nominal scale. As COEFF solely belongs to SEE models and *vice versa*, COEFF is omitted from the regression due to exact collinearity. We discuss the impact of this anomaly when we turn to meta regression results.

As pointed out, multiplier calculations also differ concerning the time horizon of measurement (Brückner and Tuladhar 2010: 16), so we list the number of quarters after the shock (HORIZON) on which the multiplier calculation is based. By collecting both the calculation method and the horizon, we can account for the effect that peak multipliers are usually recorded on a shorter horizon than integral multipliers. Thus, the pure method specific effect is separated from the timing effect. Moreover by this combination, impact multipliers, simply fall into the category of integral multipliers with horizon 1.

Some models largely rely on calibrated parameters, while others largely base their parameter values on estimations. A potential bias with respect to this distinction of methods is controlled for by a dummy (CALIB, ESTIM) for models that are calibrated or estimated to a large part. The difference only applies to DSGE-NK and RBC models, while the other model classes are estimated by nature.

Another issue that should be controlled for is the leakage of fiscal impulses through the import channel as a country-specific effect. Using the World Bank World Development

[1] We do not distinguish the various types of taxation. Moreover, some included papers deal with multipliers from tax increases. They are treated symmetrically to multipliers from lowered taxes.

Indicators data set, we recorded the average import quota (M/GDP) of the time series and country (or group of countries) that the reported multiplier relates to. With respect to calibrated models that are not based on a certain time series, we referred to the whole available time series of the country(-group) to which the model is calibrated.

Concerning subsamples, more detailed characteristics can be taken into consideration. We build five subsamples with respect to model classes. Subsample #I, comprising DSGE-NK, RBC and MACRO models, distinguishes the characteristics mentioned above, and additionally looks for agent behavior, the modeling of the interest rate reaction, and whether the model is an open-economy model. The very same characteristics are taken into account for subsample #II that focuses on DSGE-NK and RBC models only. As for agent behavior, we record the share of Keynesian agents (KEYNES), for whom Ricardian equivalence is broken. The higher the share of Keynesian agents, the higher should be the reported multiplier. MACRO models are supposed to have a share of Keynesian agents of 100 percent in the short run.

The modeling of the interest rate can take one of four values on a nominal scale (LOANABLE, INFLATION, FIXED, ZLB), namely, on the basis of a loanable funds market, an inflation target central bank reaction function, including Taylor rules, a fixed real interest rate, and a zero lower bound setting with a fixed nominal interest rate for the central bank, where expansionary fiscal policy may lower the expected real rate of interest via a Fisher effect. Fixed real rates of interest or a ZLB regime should come with higher multipliers than the other two regimes, where crowding out via interest rates is more likely. In order to control for the disparity of open-economy models and closed-economy models, we use a dummy variable (OPEN, CLOSED). We expect closed-economy models to report higher multipliers.

Subsample #III, as a complement to subsample #II, contains all observations from MACRO, VAR and SEE approaches. Due to diversity of the model classes, there are no additional variables compared to total sample. However, subsample #IV, which is the complement to subsample #I, only includes the related VAR and SEE approaches, and thus allows for more characteristics. We record some properties of the time series that the studies draw upon. We included a normalised value of the last year of the respective time series (END), the length of the time series measured in years (LENGTH) and a dummy for annual vs. quarterly data (ANNUAL, QUARTER). LENGTH and QUARTER could be proxies for quality of the study since they contain information on the sample size. END could provide information whether more recent time series tend to have lower multipliers, as discussed in van Brusselen (2009); Bilbiie et al. (2008); Bénassy-Quéré and Cimadomo (2006); Perotti (2005).

Subsample #V applies the same characteristics as subsample #IV, but refers to VAR models only. As pointed out, there is a specific discussion in the VAR literature on identification strategies of discretionary fiscal impulses, and there are five established approaches (war episodes, narrative record, recursive VAR, structural VAR and sign restricted VAR). As with model classes for the total sample, we record the various approaches on a nominal scale with dummies (WAREPI, NARRATIVE, RECURSIVE, STRUCTURAL, SIGNRES) where each observation belongs to exactly one approach. A list of all variables can be found in Table 2.

Table 1: Descriptive statistics of reported multiplier values for model classes and fiscal impulses

	TOTAL	DSGE-NK	RBC	MACRO	SEE	VAR
N	743	278	54	94	62	255
Mean	0.8	0.8	0.5	1.0	0.6	0.9
Median	0.8	0.8	0.5	1.0	0.6	0.8
Std. Dev.	0.8	0.7	0.8	0.5	0.7	0.9
Jarque-Bera p	0.00	0.00	0.55	0.03	0.00	0.00
	GSPEND	TRANS	TAX	EMPLOY		
N	525	58	147	13		
Mean	1.0	0.4	0.6	0.8		
Median	1.0	0.3	0.5	0.9		
Std. Dev.	0.8	0.5	0.7	0.9		
Jarque-Bera p	0.00	0.00	0.00	0.36		
	SPEND	CONS	INVEST	MILIT		
N	319	95	86	25		
Mean	0.9	0.9	1.2	1.1		
Median	0.9	1.0	1.1	0.9		
Std. Dev.	0.8	0.6	0.8	0.8		
Jarque-Bera p	0.00	0.12	0.00	0.01		

4 Descriptive Statistics

This section provides a short overview on reported multiplier values. We corrected for some outliers that would have otherwise distorted the distribution. As the mean of reported multipliers is around 0.8, we excluded all observations outside the interval $[-2.2; 4]$, which is about $\mu \pm 3\sigma$. Six observations where dropped from the total sample—one on the lower end and five on the upper end of the distribution.

Table 1 reports mean and median values of reported multipliers with respect to model classes and kinds of fiscal impulse. From this mono-characteristic view, multiplier values vary widely among model classes and fiscal impulses (first rows of the table). Macro models seem to report highest multipliers, while those from RBC models seem to be lowest. Means are in a range of 0.5 to 1.0, however, one should be aware that they comprise all kinds of fiscal impulses. This is important because impulses themselves also seem to come along with very different multipliers (mid-rows). The means of reported multipliers from general public spending and public employment seem to be approximately twice as high as those from tax cuts and transfers. Splitting the group of general spending into public consumption, investment, military spending and nonspecific spending is suggestive of higher multipliers for public investment (bottom rows).

It should be stressed again that these statistics should not be (mis-)interpreted as true multiplier values, even if they stem from a comprehensive literature survey. Multiplier

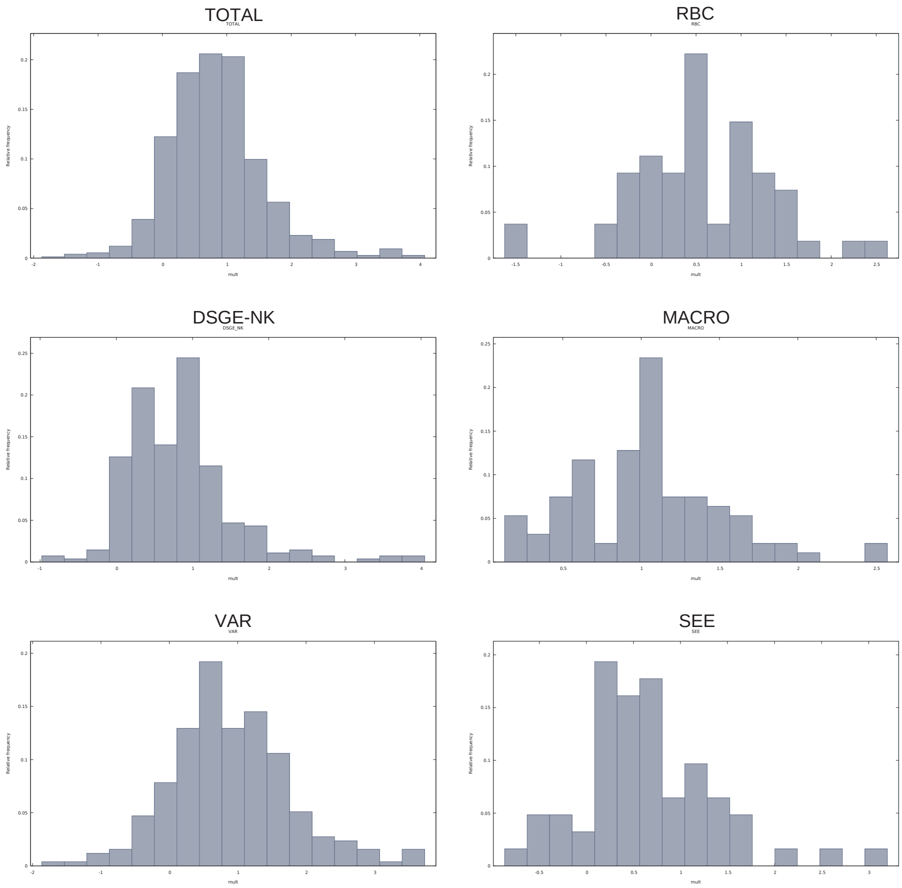

Figure 1: Histograms of reported multiplier values for various model classes

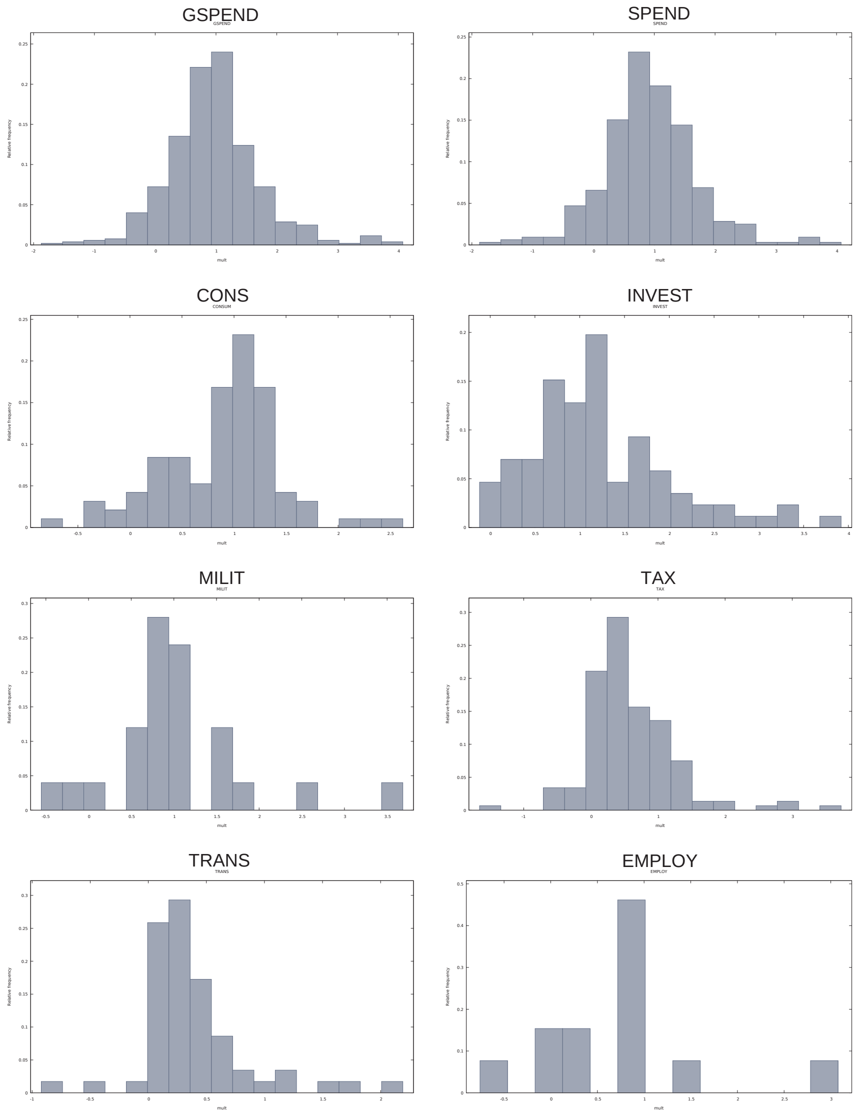

Figure 2: Histograms of reported multiplier values for various fiscal impulses

calculations may all be biased in some direction and several significant influences are unaccounted for at this point. Properties of the distribution should advise caution as well. Even though means are relatively close to medians for each model class and impulse, Figures 1 and 2 show that multipliers for the subgroups are by and large not normally distributed, which is confirmed by Jarque-Bera probabilities. Multimodal distributions point to additional influential factors. Of course, obvious distortional factors are model classes interfering the distributions of fiscal impulses and *vice versa*, but also the other variables introduced in the former section should be tested. This is why we perform a meta analysis on our sample. The aim is to separate the influences of model classes and types of impulses and to check for additional significant influences.

5 Meta Regression Analysis – Method

For the proposed meta analysis we stick to Stanley and Jarrell (2005: 302). In general, our model reads

$$k_j = \kappa + Z_j\alpha + \delta_i + \varepsilon_j \qquad j = 1, ..., N \qquad i = 1, ..., M \tag{4}$$

with

- k_j multiplier value of observation j
- κ "underlying" or "reference" multiplier value
- Z_j vector of characteristics ("moderator variables") of observation j
- α vector of systematic effects of Z_j on k_j
- δ_i vector of paper-specific intercepts (paper dummies)

We include dummies δ_i for each paper in order to control for paper-specific intercepts and use heteroscedasticity-robust estimators. However, to keep track of the main results we do not display the 89 paper dummies in regression tables (results are available on request), but we discuss their influence. Some further methodical questions need to be addressed. Meta studies often use normalisation tools to construct the effect size. To end up with a dimensionless scale, the average outcome of a treatment group is subtracted by the average outcome of the control group divided by standard deviation of the control group (Stanley 2001: 135). Normalisation is not an issue for our purpose, because the multiplier is already dimensionless. On the other hand, as mentioned above, multiplier values are not measured in a standardised manner. We control for the multiplier calculation method and the time horizon to extract comparable multiplier values, but it should be pointed out that this may only be a second best solution. However, there is no established method to translate, for example, peak multipliers into integral multipliers, or a multiplier for a horizon of ten quarters into a multiplier for five quarters.

According to Goldfarb and Stekler (2002), a general problem is double counting when several meta studies use the same data base (for instance US quarterly data from 1970-2005). Meta analysis should include only distinct and separate observations and not clones or reiterations of existing studies. However, for our purpose the same data set does not imply the same study setup. One data set can be used with different methods and model classes. These different approaches help to discriminate between specifications and should thus be included entirely.

A different question is whether to include multiple observations from one study, e. g. when the authors deal with various models, countries or types of fiscal impulse. Stanley (2001: 138) suggests to use only one observation per study or to take the average in order to control for undue weight of a single study. While this is a reasonable claim, there are some important counter-arguments. First, there is a clear trade-off with variability and degrees of freedom. Second, when picking only one observation per study, the meta analyser must take a tough decision, which one to include. Third, taking the average value may be possible for the reported multipliers, yet this technique is not valid for study characteristics of a nominal (categorical) scale type, such as the type of fiscal impulse. Fourth, taking only one observation from a comprehensive study may likewise give an undue weight to less-comprehensive studies. We and also other authors (De Grauwe and Costa Storti 2004; Nijkamp and Poot 2004; Card et al. 2010; Rusnák et al. 2011) therefore prefer including more than one observation per study. This method has been tested superior to picking only single observations per paper (Bijmolt and Pieters 2001). By using dummies for each paper, the specialty of a study is controlled for to a certain degree.

Nevertheless, we are aware of the problem of over-weighing, and thus check the robustness of our results in several ways. First, we exclude single papers with many observations ($N \geq 30$) from our sample. Second, for the total sample, we perform a robustness check by taking only one observation per study into account, namely the median value. Third, following (Sethuraman 1995), for each (sub-)sample we set up a weighted sample, by weighting each observation of a paper by the number of observations in the paper; that is, given a paper reports five different multiplier estimates, we include every estimate weighted by 1/5. The same technique applies to the study characteristics. In total every study is equally weighted. By doing so, we strike a balance between proportional influence of single studies versus degrees of freedom and variability in our survey. The resulting estimates for α are weighted least squares estimators. Results of these robustness checks can be found in the appendix.

Other meta studies differentiate the quality of included studies. Stanley (1998), for example, checks for quality on the basis of degrees of freedom, number of robustness tests and thus the number of different specifications of an included study; a higher number of degrees of freedom and different specifications should hint at a better diagnosis. De Grauwe and Costa Storti (2004) use the sample size as a quality weight. We do not perform quality selections for our total sample because the above mentioned criteria are not suitable for model-based approaches. However, for the subsamples on VAR and SEE the length and the frequency of the time series could provide information on quality.

A usual exercise for meta analyses is to control for a possible publication bias, i. e. the preference for statistically significant results in publication selection (Stanley 2008). We do not perform such a test for several reasons. First, our data set does not allow for state of the art tests of publication bias via funnel asymmetry testing or meta significance testing because RBC, DSGE and MACRO models do not produce standard errors of their results, which are necessary to apply these publication bias tests. Second, we could easily apply the more straightforward approach of Card et al. (2010), who simply set up a dummy for published versus unpublished studies. We are not convinced of this method for our purpose because we could only distinguish between journal publications (published) and working papers (unpublished). This distinction, however, is not clear cut, since major working paper series employ a refereeing process. Moreover, working papers in general are close to publication and their results have often undergone an internal publication selection by the authors. Third and most striking, after scanning the literature we do not expect a systematic preference for significant positive or negative multipliers. The aim of most of the studies is simply to calculate multiplier values, irrespective of their significance levels against zero.

6 Meta Regression Analysis – Results

In this section we present and discuss our results. For convenience, Table 2 lists the characteristics we tested for.

First, we regress reported multipliers of the total sample on characteristics as shown in Table 3. Groups of variables measured on a nominal scale, such as model class or type of impulse, are necessarily multicollinear because any observation must belong to exactly one value in this group. That is why one variable of a closed group is omitted. The influence of these omitted variables is reflected in the constant (κ), which is thus called *reference value*. It now becomes clear that κ should not be interpreted as the *true multiplier* because it depends on the reference specification.

Since we use dummies for each paper, again one of the dummies is omitted due to exact collinearity and its influence on the dependent variable is thus reflected in κ. In order to avoid a bias of the paper dummies on κ, we ran two stages of each regression. In a first step, we included all paper dummies, let the econometric software randomly choose the paper dummy to drop and calculated the mean coefficient of paper dummies. In a second step, we manually dropped the dummy closest to this mean and therefore get a reference value with a minimized bias from paper dummies. It should be pointed out that the choice of the omitted paper dummy in no way influences any of the other characteristics, but only shifts the reference value κ.

The reference for the prime estimation in column (1) is an average multiplier value calculated as an integral response to an unspecified public spending impulse, stemming from a largely estimated RBC model. Such an observation on average reports a multiplier of close to zero when controlling for other influences, which is not significantly different from zero.

Fiscal impulses differ significantly concerning their influence on the multiplier. Especially public investment produces higher multiplier values in our data set, while tax cuts and transfers have a significantly lower impact compared to direct public spending. For military spending and public employment there is only an insignificant difference to unspecified public spending.

Table 2: List of variables for meta regression

group	variable	explanation	scale
fiscal impulse			
	SPEND	unspecified public spending	dummy
	CONS	public consumption	dummy
	INVEST	public investment	dummy
	MILIT	public military spending	dummy
	GSPEND	SPEND+CONS+INVEST+MILIT	dummy
	TRANS	transfers to private sector	dummy
	EMPLOY	direct public employment	dummy
model class			
	RBC	RBC model	dummy
	DSGE-NK	New Keynesian DSGE model	dummy
	MACRO	structural macroeconometric model	dummy
	VAR	VAR model	dummy
	SEE	single equation estimation approach	dummy
VAR identification strategy			
	WAREPI	war episodes approach	dummy
	NARRATIVE	narrative record approach	dummy
	RECURSIVE	recursive VAR approach	dummy
	STRUCTURAL	structural VAR approach	dummy
	SIGNRES	sign restricted VAR approach	dummy
calibration or estimation			
	ESTIM	model more estimated than calibrated	dummy
	CALIB	model more calibrated than estimated	dummy
multiplier calculation method			
	PEAK	calculated as peak multiplier	dummy
	INTEGRAL	calculated as integral multiplier	dummy
	COEFF	calculated from coefficient of fiscal impulse	dummy
	HORIZON	horizon of the multiplier calculation	quarters after shock
open economy leakage			
	M/GDP	import quota of the surveyed country sample	percentage points
	OPEN	open economy model	dummy
	CLOSED	closed economy model	dummy
share of Ricardian vs. Keynesian agents			
	KEYNES	share of Keynesian agents	percentage points
modeling of interest rate reaction			
	LOANABLE	loanable funds market	dummy
	INFLATION	fixed real interest rate	dummy
	ZLB	zero lower bound / fixed nominal interest rate	dummy
properties of time series			
	END	normalized end of the series	percentage points
	LENGTH	length of the series	years
	QUARTER	quarterly data	dummy
	ANNUAL	annual data	dummy

Table 3: Total sample (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain[b]	(3) macro-ref[c]	(4) dsge-nk-ref[d]	(5) gspend-ref[e]	(6) no-see[a]
κ	−0.009089	0.3147	1.173***	0.7571**	0.1445	−0.07715
	(0.3121)	(0.3051)	(0.3308)	(0.3131)	(0.3255)	(0.7533)
fiscal impulse						
CONS	0.2655**	0.2682**	0.2655**	0.2655**		0.2873**
	(0.1157)	(0.1194)	(0.1157)	(0.1157)		(0.1179)
INVEST	0.5843***	0.5485***	0.5843***	0.5843***		0.6143***
	(0.1260)	(0.1290)	(0.1260)	(0.1260)		(0.1315)
MILIT	−0.1898	−0.2196	−0.1898	−0.1898		−0.1021
	(0.3168)	(0.3237)	(0.3168)	(0.3168)		(0.5340)
TAX	−0.3019***	−0.3086***	−0.3019***	−0.3019***	−0.4562***	−0.2944***
	(0.08131)	(0.08438)	(0.08131)	(0.08131)	(0.06949)	(0.08289)
TRANS	−0.3468***	−0.3465***	−0.3468***	−0.3468***	−0.6240***	−0.3415***
	(0.09694)	(0.09810)	(0.09694)	(0.09694)	(0.07597)	(0.1029)
EMPLOY	0.2221	0.2130	0.2221	0.2221	−0.03012	0.2921
	(0.2534)	(0.2708)	(0.2534)	(0.2534)	(0.2373)	(0.2671)
model class						
RBC			−1.182***	−0.7662***		
			(0.2484)	(0.2327)		
DSGE-NK	0.7662***	0.6983***	−0.4159***		0.7904***	0.8178***
	(0.2327)	(0.2324)	(0.1067)		(0.2357)	(0.2799)
MACRO	1.182***	1.142***		0.4159***	1.197***	1.249***
	(0.2484)	(0.2428)		(0.1067)	(0.2472)	(0.3009)
VAR	0.8154***	0.6420***	−0.3667	0.04920	0.7796***	0.9066***
	(0.2591)	(0.2411)	(0.2778)	(0.2650)	(0.2544)	(0.3063)
SEE	0.9393***	0.3112	−0.2428	0.1731	0.8603***	
	(0.2442)	(0.1913)	(0.2624)	(0.2478)	(0.2421)	
control variables						
CALIB	0.2156*		0.2156*	0.2156*	0.2062*	0.2080*
	(0.1134)		(0.1134)	(0.1134)	(0.1138)	(0.1150)
PEAK	0.4377***		0.4377***	0.4377***	0.3995***	0.4461***
	(0.1162)		(0.1162)	(0.1162)	(0.1165)	(0.1171)
HORIZON	0.01747***		0.01747***	0.01747***	0.01515**	0.01779***
	(0.006213)		(0.006213)	(0.006213)	(0.006013)	(0.006292)
M/GDP	−1.328***		−1.328***	−1.328***	−1.335***	−1.320***
	(0.3222)		(0.3222)	(0.3222)	(0.3277)	(0.3265)
N	743	743	743	743	743	681
Adj.R^2	0.3707	0.3389	0.3707	0.3707	0.3404	0.3527
ℓ	−615.7	−636.4	−615.7	−615.7	−634.9	−573.4

[a] reference: SPEND, RBC, ESTIM, INTEGRAL
[b] reference: SPEND, RBC
[c] reference: SPEND, MACRO, ESTIM, INTEGRAL
[d] reference: SPEND, DSGE-NK, ESTIM, INTEGRAL
[e] reference: GSPEND, RBC, ESTIM, INTEGRAL
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

The next rows show the influences of other model classes, which are all significantly higher than for the RBC specification. Of the four alternative model classes, DSGE-NK models tend to report the lowest multipliers, while MACRO models are on the upper end of the scale.

Attention should be given to control variables. Calibrated models tend to report higher multipliers than estimated ones. Estimated models were chosen as reference in order to better compare the estimated variants of DSGE and RBC models to MACRO, VAR and SEE approaches, which are estimated by nature. Peak multipliers are, as expected, significantly higher than integral multipliers. A longer horizon of measurement comes along with higher multipliers. Import quotas are highly significant with a negative impact on reported multipliers.

To do some first robustness checks, we estimated some variants of the regression in column (1). Column (2) shows a plain model without control variables. Results of our prime model are reconfirmed by and large. However, excluding control variables increases the reference multiplier and renders the difference between RBC and SEE approaches insignificant. A stepwise exclusion of controls shows that both effects largely depend on the inclusion of PEAK. This is reasonable given that peak multipliers are significantly higher and that single equation estimations cannot be differentiated with respect to peak or integral multipliers. Not controlling for this difference shifts the reference multiplier upwards and SEE multipliers downwards relative to all other model classes. The stepwise exclusion of control variables does not affect the other controls except for HORIZON that becomes insignificant when excluding PEAK. This effect is coherent with our reasoning that peak multipliers are usually recorded on a shorter horizon than integral multipliers. Ignoring the heterogeneity of peak and integral multipliers obscures the specific information of HORIZON.

The regression model in column (3) tests the impact of exchanging the reference model class. Using observations from MACRO models as reference only affects the constant and the model class group. The test reveals that DSGE-NK and RBC models report significantly lower multipliers, while VAR and SEE do not. When DSGE-NK models serve as reference, as shown in column (4), VAR and SEE coefficients are also insignificant, while the coefficent for RBC is significantly lower and the one for MACRO models is significantly higher. Taking all these information together, DSGE-NK and MACRO models are close to the more data based VAR and SEE approaches, while RBC models negatively stand out of the model classes tested.

Column (5) shows that our results are robust to a different reference fiscal impulse (GSPEND), where we do not distinguish public spending, consumption, investment and military spending. Coefficients and significance levels are only altered very slightly in comparison to column (1).

Since SEE multipliers are calculated in a different way compared to those of other model classes, in column (6) we test the robustness of our results when excluding observations from SEE models. There is almost no change in coefficients from this exercise.

The total sample comprised both model based approaches and empirical approaches, which could be criticized for comparing apples with oranges. However, both methods are coequally used for policy consulting and the meta regressions on the total sample mark their potential differences.

We now turn to subsamples in order to perform some additional robustness tests and to control for characteristics that apply to specific model classes only. We start with subsample #I, comprising observations from model based approaches (RBC, DSGE-NK, MACRO). Regression results are shown in Table 4, which provides the prime regression for this subsample as well as some simple robustness checks akin to the ones in Table 3.

Most results of the total sample are reaffirmed with subsample #I. The reference value is not significantly different from zero. Concerning fiscal impulses, public investment still significantly increases the reported multiplier, while tax cuts and transfers decrease it. Unlike the total sample, in subsample #I military spending induces the strongest multiplier effects.

Model classes differ significantly with highest value for MACRO models. Exchanging the reference model class, as done in column (4), where a DSGE-NK models serves as reference instead of a RBC model, does not alter the results for other characteristics, but reveals the difference between DSGE-NK and MACRO models insignificant for this specification.

The additional characteristics concerning agent behavior, interest rate reaction and openness to trade are all significant. As expected, the higher the share of Keynesian agents, the higher the reported multiplier. Comparing columns (1) and (2) shows that dropping additional characteristics, and thus testing the same characteristics as for the total sample, restores the significant distance between MACRO and DSGE-NK multipliers, which appeared in the total sample. Stepwise exclusion of additional characteristics reveals that agent behavior largely accounts for this difference. In other words, New Keynesian DSGE models without Ricardian agents produce insignificantly different multiplier effects as compared to structural macro models where Ricardian Equivalence does not apply in general.

Models with a loanable funds specification of the interest rate, which is our reference here, tend to have the lowest multipliers. Including a central bank reaction function with an inflation target significantly increases the multiplier. This is pretty much the same for models with a fixed real rate of interest. The highest multipliers result from models with a zero lower bound specification. When a model with a fixed real rate of interest serves as reference (column (6)), it can be shown that models with inflation targeting do not significantly differ, while the ZLB specification is still significantly higher. Other regression coefficients are unaffected by this modification. Open-economy models point to lower multipliers than closed-economy models.

The other control variables have the same algebraic sign as compared to the total sample. However, they are not significant, except for the import quota. Setting up plain regressions without control variables and additional characteristics (columns (3) and (4)) does not alter the results qualitatively.

Meta regression results for subsample #II, which is akin to subsample #I, but focuses on RBC and DSGE-NK models only, are displayed in Table 5. Subsample #II is actually a mere robustness check to #I because the same characteristics are tested. There are only some slight distinctions: it is no use to test for other reference models since there are only two possible model classes. Concerning the interest rate reaction function,

Table 4: subsample #I (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain1[b]	(3) plain2[c]	(4) plain3[d]	(5) dsge-ref[e]	(6) intfix-ref[f]
κ	−0.04350	−0.01750	0.1996	0.2873	0.6714**	0.3516
	(0.3462)	(0.2760)	(0.1778)	(0.1797)	(0.2874)	(0.3613)
fiscal impulse						
CONS	0.07406	0.1180	0.06691	0.1104	0.07406	0.07406
	(0.1011)	(0.09981)	(0.1023)	(0.1009)	(0.1011)	(0.1011)
INVEST	0.2522**	0.2542**	0.2337*	0.2360*	0.2522**	0.2522**
	(0.1220)	(0.1257)	(0.1234)	(0.1270)	(0.1220)	(0.1220)
MILIT	1.003***	0.9064***	0.8954***	0.8077***	1.003***	1.003***
	(0.2277)	(0.2319)	(0.1828)	(0.1846)	(0.2277)	(0.2277)
TAX	−0.5324***	−0.5102***	−0.5436***	−0.5204***	−0.5324***	−0.5324***
	(0.07616)	(0.08027)	(0.07820)	(0.08157)	(0.07616)	(0.07616)
TRANS	−0.6024***	−0.5986***	−0.6087***	−0.6042***	−0.6024***	−0.6024***
	(0.1047)	(0.1071)	(0.1061)	(0.1080)	(0.1047)	(0.1047)
EMPLOY	−0.004832	−0.04930	−0.01681	−0.05924	−0.004832	−0.004832
	(0.1602)	(0.1508)	(0.1680)	(0.1585)	(0.1602)	(0.1602)
model class						
RBC					−0.7149***	
					(0.2032)	
DSGE-NK	0.7149***	0.8833***	0.7214***	0.8833***		0.7149***
	(0.2032)	(0.1975)	(0.1926)	(0.1898)		(0.2032)
MACRO	0.8086**	1.349***	0.8271***	1.353***	0.09367	0.8086**
	(0.3249)	(0.2158)	(0.3182)	(0.2078)	(0.2478)	(0.3249)
additional characteristics						
KEYNES	0.6544**		0.6493**		0.6544**	0.6544**
	(0.3184)		(0.3141)		(0.3184)	(0.3184)
LOANABLE						−0.3951***
						(0.1036)
INFLAT	0.3173***		0.2855***		0.3173***	−0.07782
	(0.09164)		(0.08683)		(0.09164)	(0.1044)
FIXED	0.3951***		0.3190***		0.3951***	
	(0.1036)		(0.1175)		(0.1036)	
ZLB	0.8400***		0.8010***		0.8400***	0.4449***
	(0.1196)		(0.1150)		(0.1196)	(0.1305)
OPEN	−0.3299*		−0.3298**		−0.3299*	−0.3299*
	(0.1805)		(0.1617)		(0.1805)	(0.1805)
control variables						
CALIB	0.1421	0.1611			0.1421	0.1421
	(0.1101)	(0.1131)			(0.1101)	(0.1101)
PEAK	0.2113	0.1859			0.2113	0.2113
	(0.1518)	(0.1668)			(0.1518)	(0.1518)
HORIZON	0.002891	0.002236			0.002891	0.002891
	(0.005891)	(0.006317)			(0.005891)	(0.005891)
M/GDP	−0.7339***	−0.4437**			−0.7339***	−0.7339***
	(0.2638)	(0.2179)			(0.2638)	(0.2638)
N	426	426	426	426	426	426
Adj.R^2	0.5396	0.4537	0.5299	0.4483	0.5396	0.5396
ℓ	−226.7	−266.1	−233.5	−270.5	−226.7	−226.7

[a] reference: SPEND, RBC, ESTIM, INTEGRAL, LOANABLE, CLOSED
[b] reference: SPEND, RBC, LOANABLE, CLOSED
[c] reference: SPEND, RBC, ESTIM, INTEGRAL
[d] reference: SPEND, RBC
[e] reference: SPEND, DSGE-NK, ESTIM, INTEGRAL, LOANABLE, CLOSED
[f] reference: GSPEND, RBC, ESTIM, INTEGRAL, FIXED, CLOSED
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

there is no observation with a fixed real rate of interest for this sample, which is why the variable is dropped. Notwithstanding these slight distinctions, Table 5 by and large reproduces the results from Table 4, except for the inflation target specification, which is now insignificantly different from a loanable funds specification. However, models with a zero lower bound setting still produce significantly higher multipliers. The high effect of military spending compared to spending in general seems to be a special characteristic of model based approaches, since this is not confirmed by the total sample and the following subsamples.

With subsample #III, which takes into account observations from MACRO, VAR and SEE approaches, the complement to subsample #II is tested. Results are shown in Table 6. Due to heterogeneity of the model classes, we simply test for the broad characteristics of the total sample, omitting the calibration vs. estimation dummy that does not apply for this group because all of them are estimated. Regression results in all four columns are qualitatively similar to those of the total sample. Public investment seems to be the most effective fiscal impulse in our sample. Indirect impulses, such as taxes and transfers, seem to be less effective, although results are less significant, probably due to the smaller sample size. Military spending turns insignificant again. The three model classes do not produce significantly different multipliers.

Control variables show the same pattern as in Table 3. The high significance levels of multiplier calculation method and horizon that appeared in the total sample seem to have their roots in VAR and SEE approaches, since they do not appear in subsamples #I and #II. This is perfectly in line with intuition because the IRFs of the more model-based approaches are much smoother.

Subsample #IV focuses on VAR and SEE approaches only. It is thus the complement to subsample #I. Results are displayed in Table 7. They largely affirm the previous results. Public investment seems to be most effective among fiscal impulses; tax cuts seem to be less effective in comparison to public spending, but differences are not significant; the two model classes do not make a significant difference; control variables show the familiar pattern of the total sample and subsample #III. However, excluding them in column (2) ends up with an insignificant value for κ, close to the one of the prime specification of subsample #III that tests the very same characteristics.

What can be said about the additional characteristics, which deal with the properties of the time series the included studies draw upon? Including additional characteristics shifts the reference value, as can be seen by a comparison of columns (1) and (3) with (2) and (4). That is, controlling for the properties of the time series increases the reported multiplier. When the time series ends later, the multiplier decreases significantly. This is in line with findings in the literature on the decline of the multiplier over the last six decades (van Brusselen 2009; Bilbiie et al. 2008; Bénassy-Quéré and Cimadomo 2006; Perotti 2005). One should, however, be aware that even the most recent time series in our sample do not cover a reasonable part of the effects of the stimulus packages in response to the Great Recession. The coefficient of LENGTH indicates that the longer the time series, the higher is the multiplier. Moreover, using annual instead of quarterly data tends to reduce the multiplier. Combining both information, one could draw the conclusion that studies with a larger sample size point to higher multipliers.

Table 5: subsample #II (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain1[b]	(3) plain2[c]	(4) plain3[d]	(5) int-inflat[e]
κ	0.09881	0.05715	0.2590*	0.2445	0.1494
	(0.2416)	(0.2809)	(0.1493)	(0.2212)	(0.2649)
fiscal impulse					
CONS	0.2254*	0.2700**	0.2070	0.2548*	0.2254*
	(0.1318)	(0.1275)	(0.1355)	(0.1321)	(0.1318)
INVEST	0.3517**	0.3459**	0.3184**	0.3168**	0.3517**
	(0.1514)	(0.1552)	(0.1565)	(0.1599)	(0.1514)
MILIT	0.8826***	0.8769***	0.8360***	0.8077***	0.8826***
	(0.2166)	(0.2471)	(0.1552)	(0.1853)	(0.2166)
TAX	−0.5624***	−0.5368***	−0.5837***	−0.5557***	−0.5624***
	(0.1021)	(0.1034)	(0.1074)	(0.1083)	(0.1021)
TRANS	−0.5626***	−0.5610***	−0.5811***	−0.5755***	−0.5626***
	(0.1257)	(0.1271)	(0.1308)	(0.1319)	(0.1257)
EMPLOY	0.05129	0.004835	0.02801	−0.01318	0.05129
	(0.1843)	(0.1738)	(0.1939)	(0.1836)	(0.1843)
model class					
DSGE-NK	0.7612***	0.8833***	0.7615***	0.8833***	0.7612***
	(0.1974)	(0.1982)	(0.1887)	(0.1905)	(0.1974)
additional characteristics					
KEYNES	0.6335*		0.6314**		0.6335*
	(0.3237)		(0.3185)		(0.3237)
LOANABLE					−0.05055
					(0.1342)
INFLAT	0.05055		0.05511		
	(0.1342)		(0.1349)		
ZLB	0.5786***		0.5739***		0.5281***
	(0.1349)		(0.1352)		(0.08240)
OPEN	−0.3367*		−0.3461**		−0.3367*
	(0.1796)		(0.1580)		(0.1796)
control variables					
CALIB	0.1419	0.1629			0.1419
	(0.1102)	(0.1135)			(0.1102)
PEAK	0.2430	0.2041			0.2430
	(0.1557)	(0.1714)			(0.1557)
HORIZON	0.005032	0.003458			0.005032
	(0.006497)	(0.007014)			(0.006497)
M/GDP	−0.5694*	−0.3686			−0.5694*
	(0.3148)	(0.2485)			(0.3148)
N	332	332	332	332	332
Adj.R^2	0.5455	0.4515	0.5375	0.4469	0.5455
ℓ	−189.6	−223.2	−194.9	−227	−189.6

[a] reference: SPEND, RBC, ESTIM, INTEGRAL, LOANABLE, CLOSED
[b] reference: SPEND, RBC, ESTIM, INTEGRAL
[c] reference: SPEND, RBC, LOANABLE, CLOSED
[d] reference: SPEND, RBC
[e] reference: SPEND, RBC, ESTIM, INTEGRAL, INFLAT, CLOSED
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

Table 6: subsample #III (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain[b]	(3) see-ref[c]	(4) gspend-ref[d]
κ	0.6363	1.016***	1.135***	0.7375*
	(0.4554)	(0.3413)	(0.2267)	(0.4438)
fiscal impulse				
CONS	0.1194	0.1491	0.1194	
	(0.1728)	(0.1777)	(0.1728)	
INVEST	0.6525***	0.6003***	0.6525***	
	(0.1955)	(0.1941)	(0.1955)	
MILIT	−0.1174	−0.1760	−0.1174	
	(0.4278)	(0.4370)	(0.4278)	
TAX	−0.2288**	−0.2397**	−0.2288**	−0.3021***
	(0.1077)	(0.1110)	(0.1077)	(0.09759)
TRANS	−0.1388	−0.1096	−0.1388	−0.2542
	(0.1690)	(0.1570)	(0.1690)	(0.1575)
EMPLOY	0.3891	0.4126	0.3891	0.1311
	(0.9967)	(1.101)	(0.9967)	(0.9459)
model class				
MACRO	0.3156	0.1292	−0.1833	0.3504
	(0.3455)	(0.3403)	(0.1213)	(0.3107)
VAR			−0.4989	
			(0.3769)	
SEE	0.4989	−0.1708		0.5080
	(0.3769)	(0.3403)		(0.3424)
control variables				
PEAK	0.6706***		0.6706***	0.6233***
	(0.1584)		(0.1584)	(0.1572)
HORIZON	0.03121***		0.03121***	0.02761***
	(0.009974)		(0.009974)	(0.009613)
M_GDP	−1.864***		−1.864***	−1.830***
	(0.4959)		(0.4959)	(0.5021)
N	411	411	411	411
Adj.R^2	0.3375	0.2734	0.3375	0.2992
ℓ	−364.6	−385.4	−364.6	−378

[a] reference: SPEND, VAR, INTEGRAL
[b] reference: SPEND, VAR
[c] reference: SPEND, SEE, INTEGRAL
[d] reference: GSPEND, VAR, INTEGRAL
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

Table 7: subsample #IV (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain1[b]	(3) plain2[c]	(4) plain3[d]	(5) gspend-ref[e]
κ	2.157***	0.6115	2.008**	0.9133***	2.149***
	(0.7192)	(0.6684)	(0.7925)	(0.09140)	(0.7124)
fiscal impulse					
CONS	0.3129	0.3231	0.3804	0.3970	
	(0.3009)	(0.3004)	(0.3048)	(0.3044)	
INVEST	0.8920***	0.8973***	0.8410***	0.8541***	
	(0.2888)	(0.2872)	(0.2853)	(0.2839)	
MILIT	−0.02927	−0.003092	−0.1135	−0.06190	
	(0.4433)	(0.4393)	(0.4476)	(0.4452)	
TAX	−0.1467	−0.1510	−0.1656	−0.1746	−0.2190
	(0.1370)	(0.1386)	(0.1391)	(0.1426)	(0.1385)
TRANS	0.0004343	−0.002177	0.07596	0.07145	−0.02455
	(0.1052)	(0.1045)	(0.08863)	(0.08903)	(0.09958)
EMPLOY	0.6324	0.5835	0.7260	0.6292	0.1951
	(0.9302)	(0.9716)	(0.9938)	(1.113)	(0.8464)
model class					
SEE	0.5250	0.5330	−0.2097	−0.1877	0.5306
	(0.4045)	(0.3981)	(0.3568)	(0.3591)	(0.3469)
additional characteristics					
END	−2.024***		−2.166**		−1.968**
	(0.7680)		(0.8825)		(0.7637)
LENGTH	0.01295*		0.02575***		0.01385*
	(0.007801)		(0.008605)		(0.007434)
ANNUAL	−0.6684*		−0.3940		−0.6764**
	(0.3470)		(0.3063)		(0.3251)
control variables					
PEAK	0.7199***	0.7163***			0.6558***
	(0.1767)	(0.1702)			(0.1715)
HORIZON	0.03577***	0.03597***			0.03108***
	(0.01155)	(0.01133)			(0.01103)
M/GDP	−2.375***	−2.582***			−2.359***
	(0.6595)	(0.6834)			(0.6851)
N	317	317	317	317	317
$Adj.R^2$	0.3526	0.3446	0.2831	0.2638	0.3027
ℓ	−300.7	−304.5	−318.7	−324.7	−314.3

[a] reference: SPEND, VAR, INTEGRAL, QUARTER
[b] reference: SPEND, VAR, INTEGRAL
[c] reference: SPEND, VAR, QUARTER
[d] reference: SPEND, VAR
[e] reference: GSPEND, VAR, INTEGRAL, QUARTER
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

However, both LENGTH and ANNUAL could carry other information. LENGTH may also contain information on the sample period because longer time series on average reach farther back in time and therefore comprehend periods when multiplier effects were supposed to be higher. The quarterly vs. annual dummy may be a proxy for precision, but it may also contain a bias regarding the identification of discretionary fiscal impulses as discussed for example in Beetsma and Giuliodori (2011: F11). It should be pointed out that the coefficients of the other variables are not affected by including or excluding the additional variables.

We now turn to subsample #V in order to contribute to the discussion on different identification methods in the VAR literature on fiscal impulses. Results are reported in Table 8. To start with similarities to other subsamples, again, public investment seems to be the most effective fiscal impulse. Furthermore, additional characteristics and control variables are largely in line with those of subsample #IV.

The reference value κ is positive and significant throughout most columns. The significant values of κ in columns (1), (3), (5) and (6) depend on the inclusion of the additional characteristics, as the plain specifications of column (2) and (4) reveal.

As a central issue, the estimation tests the influence of the various identification strategies. The prime specification, including all additional characteristics and control variables shows insignificant differences between identification strategies. They, however, turn significant when excluding control variables in columns (3) and (4). Recursive VARs, Structural VARs and VARs using the narrative record all seem to have a similar positive impact on the reported multiplier in comparison to the sign restricted VARs and those based on war episodes. A stepwise inclusion of control variables reveals that whether differences are significant or not, depends on the inclusion of PEAK and HORIZON, and not on the additional characteristics or the import quota. The impact of PEAK and HORIZON is not surprising, since the different identification strategies come along with specific shapes of impulse response functions that are also connected to multiplier calculation method and horizon. Thus, one could conclude that a big part of the difference among identification strategies concerning reported multipliers is simply a question of timing of measurement. This result is in line with Caldara and Kamps (2008: 28); it is confirmed by column (5) because dropping identification strategies even makes a slight increase of the adjusted R^2 in comparison to the prime specification. Changing the reference identification strategy, as in column (6), shifts the reference value and the coefficients of other identification strategies, but does not alter the other results.

Before we conclude, we would like to refer to the statistical appendix that contains further robustness checks concerning a possible overweighing of comprehensive studies. The columns of Table 9 show results for the prime specification of our total sample, when dropping single papers with many observations ($N \geq 30$) from the sample. In Table 10 we test weighted versions of the prime specifications of all (sub-)samples by weighting each observation of a paper by the number of observations in the paper. Finally, Table 11 presents alternative specifications of the dependent variable. In column (1) we test a model using only the median multiplier of each study of the total sample. In column (2) and (3) we test a probit and logit model respectively, where the dependent variable is binary, signalling whether the multiplier is greater than or equal to one or whether

Table 8: subsample #V (Dep. Var.: multiplier)

	(1) prime[a]	(2) plain1[b]	(3) plain2[c]	(4) plain3[d]	(5) no-ident[e]	(6) recursive-ref[f]
κ	1.568*	0.3371	1.492*	0.4992	1.958***	1.978**
	(0.8978)	(0.4269)	(0.8275)	(0.6798)	(0.7231)	(0.9396)
fiscal impulse						
CONS	0.5319	0.5337	0.5863*	0.5870*	0.5384	0.5319
	(0.3396)	(0.3386)	(0.3432)	(0.3419)	(0.3359)	(0.3396)
INVEST	1.138***	1.136***	1.071***	1.071***	1.147***	1.138***
	(0.3294)	(0.3279)	(0.3261)	(0.3247)	(0.3260)	(0.3294)
MILIT	0.1575	0.1587	0.06835	0.06839	0.1685	0.1575
	(0.6361)	(0.6318)	(0.6559)	(0.6516)	(0.6300)	(0.6361)
TAX	−0.1447	−0.1441	−0.1857	−0.1857	−0.1259	−0.1447
	(0.1485)	(0.1502)	(0.1518)	(0.1566)	(0.1428)	(0.1485)
TRANS	−0.08965	−0.09171	0.08000	0.08000	−0.09460	−0.08965
	(0.1488)	(0.1479)	(0.08814)	(0.09360)	(0.1495)	(0.1488)
EMPLOY	1.352	1.249	1.475	1.264	1.359	1.352
	(1.141)	(1.238)	(1.188)	(1.442)	(1.132)	(1.141)
identification strategy						
RECURSIVE	0.4104	0.4089	0.5325**	0.5324**		
	(0.2633)	(0.2630)	(0.2139)	(0.2131)		
STRUCTURAL	0.4086	0.4075	0.4985**	0.4984**		−0.001857
	(0.2626)	(0.2624)	(0.2185)	(0.2177)		(0.2950)
SIGNRES	0.003446	0.002834	0.05113	0.05113		−0.4070
	(0.5799)	(0.5763)	(0.5752)	(0.5714)		(0.5692)
NARRATIVE	0.4086	0.4075	0.4985**	0.4984**		−0.001857
	(0.2626)	(0.2624)	(0.2185)	(0.2177)		(0.2950)
WAREPI						−0.4104
						(0.2633)
additional characteristics						
END	−2.148***		−2.454***		−2.148***	−2.148***
	(0.8125)		(0.9273)		(0.8049)	(0.8125)
LENGTH	0.01602		0.03358***		0.01582	0.01602
	(0.01007)		(0.01090)		(0.009982)	(0.01007)
ANNUAL	−0.7341*		−0.5832*		−0.7303*	−0.7341*
	(0.3836)		(0.3448)		(0.3844)	(0.3836)
control variables						
PEAK	0.7248***	0.7289***			0.7388***	0.7248***
	(0.1827)	(0.1741)			(0.1826)	(0.1827)
HORIZON	0.03571***	0.03615***			0.03676***	0.03571***
	(0.01168)	(0.01139)			(0.01191)	(0.01168)
M/GDP	−2.334***	−2.576***			−2.334***	−2.334***
	(0.6565)	(0.6883)			(0.6502)	(0.6565)
N	255	255	255	255	255	255
Adj.R^2	0.3223	0.3132	0.2437	0.2172	0.3288	0.3223
ℓ	−252.1	−255.6	−267.9	−274.1	−253.3	−252.1

[a] reference: SPEND, WAREPI, INTEGRAL, QUARTER
[b] reference: SPEND, WAREPI, INTEGRAL
[c] reference: SPEND, WAREPI, QUARTER
[d] reference: SPEND, WAREPI
[e] reference: SPEND, INTEGRAL, QUARTER
[f] reference: SPEND, RECURSIVE, INTEGRAL, QUARTER
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

it is less than one. The results of all these robustness checks largely affirm our prime specifications.

7 Conclusions

We tested a set of 89 studies on multiplier effects by meta regression analysis in order to provide a systematic overview of the different approaches, to derive stylized facts and to separate structural from method-specific effects. The method is not suitable to filter out an absolute value for the multiplier, however, it is able to extract relative differences between study characteristics.

We now draw a broad picture of our results from the meta analysis. First, reported multipliers depend on model classes. Controlling for additional variables reveals that RBC models come up with significantly lower multipliers than the rest of model classes. DSGE-NK models and MACRO models also report significantly different multipliers, however their implications are not significantly different from those of the more data oriented VAR and SEE approaches.

Second, direct public demand tends to have higher multipliers than tax cuts and transfers, even though the difference is not always significant. However, public investment seems to be the most effective fiscal impulse, a result, which is robust against many specifications. Military spending is preferred solely by the more model based approaches, especially DSGE-NK and RBC models. For VAR and SEE approaches, and for our total sample, multiplier effects of military spending do not differ from those of public spending in general.

Third, reported multipliers strongly depend on the method and horizon of calculating them. Peak multipliers are on average greater than integral multipliers and the longer the horizon of measurement, the higher is the multiplier, even though fiscal impulses are not permanent. Thus, a simple listing of multiplier values without additional information on how they were computed could provide a biased picture.

Fourth, longer time series and those with a higher frequency tend to imply higher multipliers in our sample. Time series that end in more recent years tend to imply lower multipliers. One should, however, be aware that even the most recent time series in our sample do not cover an adequate portion of the effects of the stimulus packages in response to the Great Recession.

Fifth, the more open the import channel of an economy, the lower seems to be the multiplier.

Sixth, in model based approaches the interest rate reaction function is a key parameter to the reported multiplier value. Multiplier effects are highest, when the central bank accommodates fiscal policy or is bound to a zero interest rate. Seventh, an increasing share of agents, for whom Ricardian equivalence is broken, significantly increases multiplier values. Controlling for agent behavior in our sample offsets the difference in reported multipliers of New Keynesian DSGE models and structural macro models. Both an accommodating monetary policy and liquidity constrained households correspond to the current macroeconomic setting which could imply a higher effectiveness of fiscal policy in times of the current crisis.

Eighth, divergent results of the various identification strategies in VAR models seem to be partly due to multiplier calculation methods and horizons of measurement that reflect different shapes of impulse response functions.

As an overall conclusion, we might state that reported multipliers very much depend on the setting and method chosen. Thus, economic policy consulting based on a certain multiplier study should lay open by how much specification influences the results. Our meta analysis may provide guidance concerning such influential specifications and their direction of influence.

References

Ardagna, S. (2001). Fiscal policy composition, public debt, and economic activity. *Public Choice 109*, 301–325.

Auerbach, A. J. and Y. Gorodnichenko (2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy* (forthcoming).

Baxter, M. and R. G. King (1993). Fiscal policy in general equilibrium. *American Economic Review 83*(3), 315–334.

Beetsma, R. and M. Giuliodori (2011). The Effects of Government Purchases Shocks: Review and Estimates for the EU*. *The Economic Journal 121*(550), F4–F32.

Bénassy-Quéré, A. and J. Cimadomo (2006). Changing Patterns of Domestic and Cross-Border Fiscal Policy Multipliers in Europe and the US. CEPII working paper 2006-24.

Bijmolt, T. H. A. and R. G. M. Pieters (2001). Meta-Analysis in Marketing when Studies Contain Multiple Measurements. *Marketing Letters 12*(2), 157–169.

Bilbiie, F. O., A. Meier, and G. J. Müller (2008). What Accounts for the Changes in U.S. Fiscal Policy Transmission? *Journal of Money, Credit, and Banking 40*(7), 1439–1469.

Blanchard, O. and R. Perotti (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *Quarterly Journal of Economics 117*(4), 1329–1368.

Bouthevillain, C., J. Caruana, C. Checherita, J. Cunha, E. Gordo, S. Haroutunian, G. Langenus, A. Hubic, B. Manzke, J. J. Pérez, and P. Tommasino (2009). Pros and Cons of various fiscal measures to stimulate the economy. Banque Centrale du Luxembourg working paper 40.

Briotti, M. G. (2005). Economic reactions to public finance: A survey of the literature. European Central Bank occasional paper series 38.

Brückner, M. and A. Tuladhar (2010). Public Investment as a Fiscal Stimulus: Evidence from Japan's Regional Spending During the 1990s. IMF Working Paper WP/10/110.

Caldara, D. and C. Kamps (2008). What are the effects of fiscal policy shocks? A VAR-based comparative analysis. European Central Bank working paper series 877.

Card, D., J. Kluve, and A. M. Weber (2010). Active Labour Market Policy Evaluations: A Meta-Analysis. *Economic Journal 120*, F452–F477.

Card, D. and A. B. Krueger (1995). Time-Series Minimum-Wage Studies: A Meta-analysis. *American Economic Review 85*(2), 238–243.

Carroll, C. (2009). Comments and Discussion on "By How Much Does GDP Rise If the Government Buys More Output?". *Brookings Papers on Economic Activity 2009*(2), 232–249.

Cwik, T. and V. Wieland (2011). Keynesian government spending multipliers and spillovers in the euro area. *Economic Policy 26*(67), 493–549.

De Grauwe, P. and C. Costa Storti (2004). The Effects of Monetary Policy: A Meta-Analysis. CESifo working paper 1224.

Fatás, A. and I. Mihov (2001). The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. CEPR Discussion Papers 2760.

Fatás, A. and I. Mihov (2009). Why Fiscal Stimulus is Likely to Work. *International Finance 12*(1), 57–73.

Feld, L. P. and J. H. Heckemeyer (2011). FDI and Taxation: A Meta-Study. *Journal of Economic Surveys 25*(2), 233–272.

Freedman, C., M. Kumhof, D. Laxton, D. Muir, and S. Mursula (2010). Global effects of fiscal stimulus during the crisis. *Journal of Monetary Economics 57*(5), 506–526.

Galí, J., J. D. López-Salido, and J. Vallés (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association 5*(1), 227–270.

Goldfarb, R. and H. Stekler (2002). Wheat from Chaff: Meta-Analysis as Quantitative Literature Review: Comment. *Journal of Economic Perspectives 16*(3), 225–226.

Hasset, K. A. (2009). Why Fiscal Stimulus is Unlikely to Work. *International Finance 12*(1), 75–91.

Hebous, S. (2011). The Effects of Discretionary Fiscal Policy on Macroeconomic Aggregates: A Reappraisal. *Journal of Economic Surveys 25*(4), 674–707.

Linnemann, L. (2006). The Effect of Government Spending on Private Consumption: A Puzzle? *Journal of Money, Credit, and Banking 38*(7), 1715–1735.

Mazraani, S. (2010). Public Expenditures in an RBC Model: A Likelihood Evaluation of Crowding-in and Crowding-out Effects. Mimeo.

Mountford, A. and H. Uhlig (2009). What are the Effects of Fiscal Policy Shocks? *Journal of Applied Econometrics 24*(6), 960–992.

Nijkamp, P. and J. Poot (2004). Meta-Analysis of the Effect of Fiscal Policies on Long-Run Growth. *European Journal of Political Economy 20*(?), 91–124.

Parker, J. A. (2011). On Measuring the Effects of Fiscal Policy in Recessions. *Journal of Economic Literature 49*(3), 703–718.

Perotti, R. (2005). Estimating the effects of fiscal policy in OECD countries. CEPR Discussion Papers 4842.

Ramey, V. A. (2011). Can Government Purchases Stimulate the Economy? *Journal of Economic Literature 49*(3), 673–685.

Ramey, V. A. and M. D. Shapiro (1998). Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy 48*(1), 145–194.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Rusnák, M., T. Havránek, and R. Horváth (2011). How to solve the Price Puzzle? A Meta-Analysis. Czech National Bank Working Paper Series 2/2011.

Sethuraman, R. (1995). A meta-analysis of national brand and store brand cross-promotional price elasticities. *Marketing Letters 6*(4), 275–286.

Spilimbergo, A., S. Symansky, and M. Schindler (2009). Fiscal Multipliers. IMF Staff Position Note SPN/09/11.

Stanley, T. D. (1998). New Wine in Old Bottles: A Meta-Analysis of Ricardian Equivalence. *Southern Economic Journal 64*(3), 713–727.

Stanley, T. D. (2001). Wheat from Chaff: Meta-Analysis as Quantitative Literature Review. *Journal of Economic Perspectives 15*(3), 131–150.

Stanley, T. D. (2008). Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection. *Oxford Bulletin of Economics and Statistics 70*(1), 103–127.

Stanley, T. D. and S. B. Jarrell (2005). Meta-regression Analysis: A Quantitative Method of Literature Surveys. *Journal of Economic Surveys 19*(3), 299–308.

van Brusselen, P. (2009). Fiscal Stabilisation Plans and the Outlook for the World Economy. ENEPRI working paper 55.

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

Statistical Appendix

The following tests check for the robustness of our results with respect to a possible overweighing of single studies with many observations and alternative specifications of the endogenous variable. The columns of Table 9 show results for the prime specification of our total sample, when dropping single papers with many observations ($N \geq 30$) from the sample. In Table 10 we test weighted versions of the prime specifications of all (sub-)samples by weighting each observation of a paper by the number of observations in the paper. Finally, Table 11 presents alternative specifications of the dependent variable. In column (1) we test a model using only the median multiplier of each study of the total sample. In column (2) and (3) we test a probit and logit model respectively, where the dependent variable is binary, signalling whether the multiplier is greater than or equal to one or whether it is less than one. The results of all these robustness checks largely affirm our prime specifications.

Table 9: Robustness of total sample – stepwise exclusion of comprehensive studies (Dep. Var.: multiplier)

	(1) no d6	(2) no d45	(3) no d46	(4) no d48	(5) no d51
κ	-0.09781	-0.007616	0.03048	0.01297	-0.006129
	(0.3263)	(0.3159)	(0.3168)	(0.2795)	(0.3569)
fiscal impulse					
CONS	0.1764	0.2649^{**}	0.2835^{**}	0.2956^{***}	0.2945^{**}
	(0.1391)	(0.1174)	(0.1178)	(0.1063)	(0.1211)
INVEST	0.6121^{***}	0.5825^{***}	0.5679^{***}	0.4871^{***}	0.6112^{***}
	(0.1533)	(0.1275)	(0.1303)	(0.1227)	(0.1304)
MILIT	-0.1911	-0.1910	-0.1911	-0.2525	-0.1742
	(0.3195)	(0.3187)	(0.3176)	(0.2563)	(0.3190)
TAX	-0.3032^{***}	-0.3026^{***}	-0.3012^{***}	-0.3045^{***}	-0.2590^{***}
	(0.08411)	(0.08342)	(0.08167)	(0.07993)	(0.09316)
TRANS	-0.3070^{***}	-0.3472^{***}	-0.3470^{***}	-0.3695^{***}	-0.3189^{***}
	(0.1133)	(0.09846)	(0.09727)	(0.09468)	(0.1028)
EMPLOY	0.2126	0.2218	0.2224	0.2016	0.2470
	(0.2510)	(0.2563)	(0.2542)	(0.2512)	(0.2535)
model class					
DSGE-NK	0.7925^{***}	0.7643^{***}	0.7635^{***}	0.7645^{***}	0.7515^{***}
	(0.2414)	(0.2330)	(0.2338)	(0.2332)	(0.2350)
MACRO	1.214^{***}	1.183^{***}	1.182^{***}	1.180^{***}	1.412^{***}
	(0.2640)	(0.2476)	(0.2492)	(0.2484)	(0.3405)
VAR	0.9987^{***}	0.8176^{***}	0.8151^{***}	0.8199^{***}	0.8171^{***}
	(0.3610)	(0.2589)	(0.2601)	(0.2596)	(0.2604)
SEE	1.143^{***}	0.9074^{***}	0.9069^{***}	0.9234^{***}	0.9523^{***}
	(0.3728)	(0.2488)	(0.2469)	(0.2472)	(0.2476)
control variables					
CALIB	0.4198	0.2156^{*}	0.2137^{*}	0.2109^{*}	0.2233^{*}
	(0.3094)	(0.1131)	(0.1130)	(0.1126)	(0.1146)
PEAK	0.4412^{***}	0.4043^{***}	0.4065^{***}	0.4157^{***}	0.4357^{***}
	(0.1173)	(0.1272)	(0.1221)	(0.1181)	(0.1164)
HORIZON	0.01805^{***}	0.01678^{***}	0.01504^{**}	0.01693^{***}	0.01725^{***}
	(0.006320)	(0.006478)	(0.006559)	(0.006398)	(0.006267)
M/GDP	-1.696^{***}	-1.175^{***}	-1.321^{***}	-1.329^{***}	-1.565^{***}
	(0.4219)	(0.3062)	(0.3365)	(0.3232)	(0.3857)
N	667	713	709	713	710
$Adj.R^2$	0.3504	0.3600	0.3737	0.3857	0.3648
ℓ	-579.2	-580.8	-587.4	-553.2	-599.7

reference: SPEND, RBC, ESTIM, INTEGRAL
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

Table 10: Robustness of samples – weighted (sub-)samples (Dep. Var.: weighted multiplier)

	(1) total-w[a]	(2) #I-w[b]	(3) #II-w[b]	(4) #III-w[c]	(5) #IV-w[d]		(6) #V-w[e]
κ	-0.01139	0.006469	0.007924	0.2015*	0.2373*		0.2674***
	(0.007391)	(0.007499)	(0.008752)	(0.1247)	(0.1522)		(0.09470)
fiscal impulse							
CONS	0.11	-0.1595	-0.06246	0.5457*	0.8268		0.4023
	(0.1836)	(0.164)	(0.2153)	(0.284)	(0.5202)		(0.5108)
INVEST	0.5521***	-0.1195	-0.2489	0.8140**	1.293**		0.5659
	(0.2092)	(0.2265)	(0.2356)	(0.3631)	(0.6172)		(0.5979)
MILIT	0.12	0.1367	0.1471	0.4402**	0.4664***		0.05593
	(0.1260)	(0.1102)	(0.1147)	(0.1762)	(0.1776)		(0.6769)
TAX	-0.12	-0.6530***	-0.6361***	0.1505	0.2457		0.2122
	(0.2116)	(0.1595)	(0.1704)	(0.2342)	(0.2355)		(0.2743)
TRANS	-0.4780***	-0.6023***	-0.7545***	-0.1359	0.1966		-0.05503
	(0.1696)	(0.1667)	(0.195)	(0.2474)	(0.1436)		(0.1889)
EMPLOY	0.41	-0.4278**	-0.3991**	0.03525	0.1607		0.995
	(0.3056)	(0.1805)	(0.1897)	(0.5279)	(0.5794)		(1.203)
model class						*identif. strategy*	
RBC	ref	ref	ref			WAREPI	ref
DSGE-NK	0.7543***	0.2291	0.2145			RECURSIVE	0.3427
	(0.1519)	(0.144)	(0.148)				(0.2464)
MACRO	1.400***	-0.2138		0.2527		STRUCTURAL	0.3371
	(0.1912)	(0.355)		(0.4154)			(0.2152)
VAR	1.226***			ref	ref	SIGNRES	-0.08239
	(0.1752)						(0.5577)
SEE	0.9518***			0.3671	0.547	NARRATIVE	0.3371
	(0.1738)			(0.5023)	(0.6076)		(0.2152)
additional characteristics							
KEYNES		0.9958***	1.056***				
		(0.282)	(0.2734)				
INFLAT		-0.05215	-0.04557				
		(0.1185)	(0.1222)				
FIXED		0.043					
		(0.17)					
ZLB		1.034***	1.107***				
		(0.1475)	(0.1502)				
OPEN		0.1009	0.05038				
		(0.08536)	(0.09711)				
END					-1.870***		-1.886**
					(0.6787)		(0.7282)
LENGTH					0.01085***		0.01520**
					(0.003984)		(0.006608)
ANNUAL					-0.6439		-0.7237
					(0.4687)		(0.4705)
control variables							
CALIB	0.6380***	0.3717***	0.3789***				
	(0.1734)	(0.09016)	(0.1019)				
PEAK	0.18*	0.5736***	0.5595***	0.5181*	0.7354*		0.6979*
	(0.1230)	(0.144)	(0.1638)	(0.3044)	(0.4146)		(0.4324)
HORIZON	0.01888***	0.02088***	0.02059***	0.01729	0.0358		0.02843
	(0.004130)	(0.003478)	(0.003851)	(0.01741)	(0.02841)		(0.02892)
M/GDP	-1.924***	-1.387***	-1.462**	-2.382***	-2.527**		-2.488***
	(0.6051)	(0.4939)	(0.599)	(0.7299)	(0.9838)		(0.8933)
N	743	426	332	411	317		255
Adj.R^2	0.7065	0.8529	0.8077	0.9037	0.8905		0.8852
ℓ	601.30	448.4	323.3	498.9	374.1		271.4

[a] reference: SPEND, RBC, ESTIM, INTEGRAL
[b] reference: SPEND, RBC, LOANABLE, ESTIM, INTEGRAL
[c] reference: SPEND, RBC, INTEGRAL
[d] reference: SPEND, RBC, INTEGRAL, QUARTER
[e] reference: SPEND, WAREPI, INTEGRAL, QUARTER
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

Table 11: Robustness of total sample – median values; probit and logit model (Dep. Var.: multiplier)

	(1) median	(2) probit[a]	(3) logit[a]
κ	0.4007	-3.2739***	-6.579***
	(0.3032)	(0.9867)	(1.881)
fiscal impulse			
CONS	-0.2059	0.8743***	1.441**
	(0.1539)	(0.3146)	(0.5985)
INVEST	0.1142	1.342***	2.197***
	(0.2715)	(0.2915)	(0.5422)
MILIT	0.1575	-0.7769	-1.342
	(0.1857)	(0.5361)	(0.9503)
TAX	-0.1279	-0.8235***	-1.502***
	(0.161)	(0.2167)	(0.4138)
TRANS	-0.6710***	-1.476***	-2.601***
	(0.1604)	(0.3339)	(0.6667)
EMPLOY	-0.3827**	0.8987	1.303
	(0.1778)	(0.6747)	(1.552)
model class			
DSGE-NK	0.5355**	2.077***	3.784***
	(0.2083)	(0.7488)	(1.347)
MACRO	1.117***	8.887***	23.84***
	(0.2386)	(0.8418)	(1.223)
VAR	0.6621***	2.130***	3.874***
	(0.2398)	(0.8055)	(1.419)
SEE	0.437	1.084	2.094
	(0.3275)	(0.995)	(2.344)
control variables			
CALIB	0.2623	0.3683	0.5849
	(0.1852)	(0.2727)	(0.4927)
PEAK	0.0196	0.8480***	1.481***
	(0.1072)	(0.2546)	(0.4706)
HORIZON	0.01549**	0.03977***	0.06721***
	(0.005936)	(0.01348)	(0.02494)
M/GDP	-1.451**	-2.698***	-4.556***
	(0.6899)	(0.9441)	(1.674)
N	89	641	641
$Adj.R^2$	0.3352	0.3062	0.3085
ℓ	-49.69	-298.9	-297.9

reference: SPEND, RBC, ESTIM, INTEGRAL
*, **, *** indicate significance at the 10, 5, 1 percent level respectively
Standard errors in parentheses

[a] dependent variable is binary value signalling, whether reported multiplier is $<\geq 1$.

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, http://www.imk-boeckler.de

IMK Working Paper is an online publication series available at:
http://www.boeckler.de/imk_5016.htm

ISSN: 1861-2199

Hans Böckler Stiftung

Fakten für eine faire Arbeitswelt.