IMK
Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

September 2013

117

Working Paper

Sebastian Gechert[1]

What fiscal policy is most effective? A meta regression analysis*

September 8, 2013

Abstract

We apply meta regression analysis to a unique data set of 104 studies on multiplier effects in order to derive stylized facts and to quantify the differing effectiveness of the composition of fiscal impulses, adjusted for study-design characteristics. We find that public spending multipliers are close to one and about 0.3 to 0.4 units larger than tax and transfr multipliers. Public investment multipliers are found to be even larger than those of spending in general by approximately 0.5 units. Reported multipliers vary with study-design, whose influence should be laid open when drawing policy conclusions. Our meta analysis provides guidance concerning influential factors, their sign and magnitude.

JEL classification: E27, E62, H30.

Keywords: multiplier effects; fiscal policy; meta analysis.

[1] Chemnitz University of Technology and Macroeconomic Policy Institute,
Hans-Boeckler-Strasse 39, 40476 Düsseldorf, Germany.
E-mail: sebastian-gechert@boeckler.de. Tel: +49 211 7778 306

* We are grateful to Lisa Hahn, 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 thank Fritz Helmedag, Tom Stanley, Till van Treek and Henner Will. All errors are ours..

What fiscal policy is most effective? A meta regression analysis*

Sebastian Gechert†

September 8, 2013

Abstract. We apply meta regression analysis to a unique data set of 104 studies on multiplier effects in order to derive stylized facts and to quantify the differing effectiveness of the composition of fiscal impulses, adjusted for study-design characteristics. We find that public spending multipliers are close to one and about 0.3 to 0.4 units larger than tax and transfer multipliers. Public investment multipliers are found to be even larger than those of spending in general by approximately 0.5 units. Reported multipliers vary with study-design, whose influence should be laid open when drawing policy conclusions. Our meta analysis provides guidance concerning influential factors, their sign and magnitude.

Keywords. multiplier effects; fiscal policy; meta analysis

JEL classification. E27, E62, H30

*We are grateful to Lisa Hahn, Rafael Wildauer and Michael Hachula for their excellent research assistance. We would like to thank the authors of the studies we analyzed for providing us with additional information. Moreover we thank Fritz Helmedag, Katja Rietzler, Tom Stanley, Till van Treeck and Henner Will. All errors are ours.

†Chemnitz University of Technology and Macroeconomic Policy Institute, Hans-Boeckler-Strasse 39, 40476 Düsseldorf, Germany. E-mail: sebastian-gechert@boeckler.de. Tel: +49 211 7778 306.

1 Introduction

The size of the fiscal multiplier has become a hot topic since the financial crisis and the subsequent Euro Area crisis. In particular, the question as to the effects of the American Recovery and Reinvestment Act (ARRA) were a source of economic discussion. Turning to the more recent austerity debate in the US and Europe, the effects of fiscal contractions on growth 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.

The empirical literature on the size of the multiplier is growing fast, tackling the issue with manifold model classes, identification strategies, and specifications. Results are far from a consensus, ranging from expansionary austerity (negative multipliers) to self-defeating austerity (large multipliers).

The vast array of different approaches and assumptions necessitates a systematic analysis of the existing evidence. 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 the reported multiplier values in the literature are quantifiable, it is possible to review the outcomes from the literature with statistical criteria and quantify the specific influence of a study characteristic. Meta regression analysis is therefore a suitable tool for this issue.

This paper is a substantial update on an earlier version (Gechert and Will 2012). We apply meta regression analysis to a unique data set of 104 studies on multiplier effects providing us with 1069 multiplier observations. Our aim is two-fold: First, we provide a systematic overview of the literature to derive stylized facts on influential factors. Second, and more specifically, we quantify the differing effectiveness of the composition of fiscal impulses, adjusted for the interference of study-design characteristics and sample specifics.

It should be stressed that our method is not suitable for finding *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. However, our approach allows us to make statements on the relative effectiveness of fiscal shocks and effects of study specifics. We test the robustness of our results in various dimensions.

Our main results are as follows: First, multipliers from public spending are significantly positive and on average close to one, yet they vary a lot with study design and the underlying sample. Second, multipliers of direct public demand impulses exceed those of tax cuts and transfers by 0.3 to 0.4 units. Public investment in our analysis turns out to be the most effective fiscal impulse with multipliers 0.3 to 0.8 units above public spending in general. Third, reported multipliers depend on model classes with RBC models yielding multipliers close to zero and backward-looking macroeconometric models reporting significantly higher multipliers than the reference VAR models. Multipliers from single equation estimations also seem to have somewhat higher multipliers when controlling for additional factors. New Keynesian DSGE models report multipliers fairly close to those of VAR models, however, this finding does not apply to large-scale estimated DSGE models whose multipliers are significantly lower. Fourth, reported multipliers from estimation-based approaches strongly depend on the method and horizon of calculating them. Peak multipliers are on average 0.3 units greater than cumulative multipliers and the longer the horizon of measurement, the higher is the multiplier. Thus, a simple listing of multiplier values without additional information on how they were computed could give a biased picture. Multipliers from simulation-based approaches are largely insensitive to these issues, which raises questions as to the power of simulations to predict the timing of the effects of fiscal policy. Fifth, the more open the import channel of an economy, the lower is the multiplier.

Sixth, when controlling for fiscal impulses and study-design characteristics there are some weak signs that more precise studies report higher multipliers. Following Stanley (2008), this could point to a negative publication selection bias, but it should be stressed that our measure of precision is only second-best and results are not robust. Seventh, the various identification strategies to deal with endogeneity of fiscal impulses in estimation-based approaches report rather different multipliers. Results from the lion's share of observations in the literature stemming from the narrative record, the Blanchard-Perotti method, the recursive and instrument-variable approaches roughly point to the same multipliers on the upper end of the scale, close to one. Using cyclically-adjusted budget variables clearly stands out negatively with multipliers close to zero. Eighth, time series from more recent years tend to yield lower multipliers, confirming the findings in 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 an adequate portion of the effects of the stimulus packages in response to the Great Recession. Ninth, setting up a crisis scenario for the average of the whole set of models in our simulation-based sample yields multipliers close to two, which implies a higher effectiveness of fiscal policy in times of the recent crisis years and a more negative impact of the concerted austerity measures in the Euro Area.

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

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 descriptive statistics. Section four explains the meta regression method and discusses methodical issues. Section five introduces the set of characteristics that serve as explanatory variables in the multivariate meta regression. Section six provides regression results, including various robustness checks. The final section concludes.

2 Literature Review

2.1 Other Meta Analyses on Macroeconomic Issues

To our knowledge, the present study and its previous version (Gechert and Will 2012) are the first application of meta regression analysis on the growing literature on fiscal multipliers. There are some similar studies on other macroeconomic policy evaluations. De Grauwe and Costa Storti (2004) meta analyze 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) analyze 97 studies on active labor market policies and evaluate the effectiveness of certain kinds of programs. The seminal meta analysis of Card and Krueger (1995) provided insights into the reported effects of minimum wages depending on the study specification. Feld and Heckemeyer (2011) meta analyze 45 studies on the tax semi-elasticity of FDI. An overview of some further meta studies in economics can be found in Stanley (2001: 134), and more recently in Stanley and Doucouliagos (2012: 39).

2.2 Other Surveys on Fiscal Multipliers

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 (2011a) surveys both the theoretical and empirical literature and works with representative examples for specific model classes, data sets and macroeconomic preconditions. She makes a rough estimate for multipliers from transitory public spending in a range that is "probably between 0.8 and 1.5". Parker (2011) focuses on the deviations from standard multiplier effects arising from nonlinearities. The study argues in favor of higher multipliers in deep recessions, but does not provide quantitative outcomes. Fatás and Mihov (2009) argue in a similar direction, drawing on selected empirical evidence and some historical examples. So do Mineshima et al. (2013), who also provide a table of average multipliers from model simulations. Hasset (2009) juxtaposes some empirical papers substantiating Keynesian effects with those underpinning neoclassical projections, and finds mixed evidence in the short run.

A survey from Hebous (2011) lists several outcomes from VAR studies, however, it only draws some general qualitative conclusions. The same applies to Bouthevillain et al. (2009) who refer to a number of empirical and simulation-based approaches. Spilimbergo et al. (2009) list a couple of studies from simulations and empirical papers. They find some rough rules of thumb for multiplier values depending on the size and openness of the surveyed country between 0.5 and 1.5. It is, however, not clear how these values are derived and how other study characteristics impact the multiplier value. The survey in van Brusselen (2009) gives a range of multiplier values between -4.8 and 3.8, derived from different model classes and distinguishing changes in taxes and spending. The author points to the lack of coherence and the sensitivity of results with respect to specifications.

While at least some of these studies include tables of study results and study characteristics to categorize the existing literature, there is a lack of a systematic statistical analysis of the quantity and significance of the influence of study specifications on the reported multiplier. To the best of our knowledge, a quantitative survey of fiscal multipliers based on statistical criteria is still missing in the literature. The present paper is intended to fill this gap.

2.3 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 eight fiscal measures, namely public consumption (label: CONS), public investment (label: INVEST), military spending (label: MILIT), direct public employment (label: EMPLOY), transfers to households (label: TRANS) and tax cuts (label: TAX), notwithstanding more detailed classifications thereof. Many studies do not distinguish between public consumption, investment and military spending, but simply refer to public spending (label: SPEND); some do not even distinguish between spending and revenue categories and simply make use of deficit spending (label: DEFICIT) without any detailing.

From a theoretical perspective, arguments of different strands in the literature would allow for ambiguous rankings of the relative effectiveness of fiscal impulses. Some studies argue in favor of direct public spending because of the full first-round impact on effective demand. Some point to the particular role of public investment because of its long-run impact on growth that could push short-run expectations as well. Some claim higher multipliers for tax reliefs or transfers due to a lower crowding-out effect as compared to direct public spending. Some point to a high impact of military expenses on growth due to its monopsonistic nature, while others expect stronger crowding-out from military spending as war times usually come along with high industrial capacity utilization.

Empirical results are very mixed and it is not easy to find *prima facie* evidence concerning the relative effectiveness of fiscal measures, particularly because study designs differ. Our hypothesis is that some of the variety of results is due to a lot of interfering study characteristics, whose particular impact we try to identify and adjust for via meta regression analysis in order to get a clearer picture of the pure impact stemming from the type of fiscal impulse.

The most basic distinction regarding study design is the model class employed. Our survey includes simulation-based studies as well as pure empirical investigations. We discriminate between New Classical RBC (or D(S)GE) models, New Keynesian DSGE models, structural macroeconometric models, VAR models, and various single equation estimation techniques (OLS, ML, GMM, ECM, ...).

Basic new classical RBC or D(S)GE models (label: RBC) entail a utility maximizing, representative household for whom Ricardian equivalence holds. Additionally, they feature fully competitive labor and goods markets. Basically, 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 labor supply and with distortional effects of taxation (Ardagna 2001; Fatás and Mihov 2001).

Most contemporary simulation-based studies on fiscal multipliers use New Keynesian DSGE models (label: DSGE-NK), extending the basic RBC model by introducing monopolistic competition and sticky prices 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 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. These 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 can yield higher multiplier values. Ricardian equivalence is broken by assuming high individual discount rates or liquidity constraints for some households. They are subsumed under the heading of *non-Ricardian 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. Models with these specifications can yield multipliers far above one.

The third type of models are structural macroeconometric models (label: MACRO), which typically do not incorporate utility-maximizing agents, but estimate backward-looking 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 central here, which usually leads to multipliers larger than one by crowding-in of private consumption or investment, depending on the monetary and foreign-trade regime.

The more empirical strand of the literature applies vector autoregressive models (label: VAR) and single equation estimations (label: SEE) with varying sets of control variables and country samples, generally producing a wider range of results than the more structural approaches. Since there are obvious endogeneity problems when it comes to estimating fiscal multipliers, the literature has developed various identification schemes.

For VAR studies, there are five established approaches of identification of exogenous fiscal shocks, two of which rely on additional historical information, and three of which try to identify exogenous fiscal shocks directly from the time series.[1] (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 in the fiscal policy literature by Romer and Romer (2010), follows a similar idea, but employs historical information such as government announcements or economic forecasts, and is not limited to military spending; Romer and Romer (2010) focus on discretionary tax changes. (3) The recursive VAR approach (Fatás and Mihov 2001) uses unfiltered time series of public spending and revenue series and a Choleski decomposition that imposes a causal ordering of the variables of the VAR with zero restrictions to the factorization matrix 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-restriction VAR approach (Mountford and Uhlig 2009) identifies exogenous fiscal shocks by imposing sign restrictions directly to the impulse-response functions of the fiscal shocks and then distinguishing them from a business cycle shock.

Our data set also includes a variety of single equation estimations. To cope with endogeneity issues, these studies either build on the war episodes or narrative approach, use cyclically-adjusted fiscal time series or instrumented variables. The set of econometric techniques comprises OLS, ML, GMM, and ECM approaches from time series or panel data. The multiplier in single equation estimations usually appears in the coefficients of the fiscal variables (and their potential lags).

Besides model classes, another study-design characteristic is the means of calculating the multiplier value. In general, multiplier values are drawn from standardized fiscal impulses (e. g. one percent of GDP or one currency unit) that allow for a dimensionless comparable input-output relation. Notwithstanding some exceptions, it is generally assumed that multiplier effects are linear in scale and sign. 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

[1] see Caldara and Kamps (2008) for a comprehensive explanation of most of the methods.

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 (IRF) of standardized fiscal policy shocks. Multipliers k are calculated either as the *peak* response of GDP with respect to the initial fiscal impulse (FI_t)

$$k = \frac{\max_n \Delta Y_{t+h}}{\Delta FI_t} \tag{1}$$

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

$$k = \frac{\sum_n \Delta Y_{t+h}}{\sum_n \Delta FI_{t+h}} \tag{2}$$

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

$$k = \frac{\Delta Y_t}{\Delta FI_t} \tag{3}$$

where $\Delta(\cdot)$ marks deviation of a scenario with a shock to the fiscal variable as against a scenario without such a shock. Similar calculations apply for multipliers from changes in taxes or transfers. Equations show that additional information concerning the horizon of measurement h is needed. The calculation method and the parameter h may have a big impact on multipliers, especially when impulse and response functions do not have the same shape. Since peaks are usually the maxima of response functions of GDP, we would expect peak multipliers to exceed cumulative multipliers. However, sharply declining fiscal impulse functions combined with long-lasting GDP responses can produce cumulative multipliers exceeding peak multipliers. Impact multipliers can be subsumed under cumulative multipliers with a horizon of $h = 1$.

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. Multipliers from single equation estimations can thus be incorporated into cumulative multipliers with the respective horizon. Some authors refer to net-present value multipliers (Mountford and Uhlig 2009). Since we are interested in short horizons and both the fiscal shock and the GDP response are discounted at the same rate, present value multipliers should not deviate much from their non-discounted counterparts, such that we do not treat them separately.

3 Data Set and Descriptive Statistics

Our data set includes estimation-based and simulation-based approaches to determine short-term output effects of discretionary fiscal policy measures. It takes into account 104 papers from 1992 to 2012, providing a sample of 1069 observations of multiplier values. A list of included papers is given in Table 7 in the appendix. The majority of papers in our sample have been published from 2007 onwards, which is certainly 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, the RePEc archive and Google Scholar, as well as established working paper series (NBER, CEPR, IMF, Fed, ECB). As a necessary precondition, papers had to 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.

We corrected for some outliers. As the mean of reported multipliers is around 0.85, we excluded all observations outside the interval $[-2.2; 4]$, which is about $\mu \pm 3\sigma$. Six observations where dropped from the sample—one on the lower end and five on the upper end of the distribution.

Table 1 provides basic statistics of reported multipliers for the total sample and subsamples with respect to fiscal impulses and model classes. From this mono-characteristic view, multiplier values vary widely among model classes and fiscal impulses, and the standard deviation of each single characteristic is wide. The means of reported multipliers from general public spending seem to be approximately twice as high as those from tax cuts and transfers. Splitting the group of general spending into non-specific public spending, public consumption, investment and military spending is suggestive of higher multipliers for public investment. With respect to model classes, macro models and VAR models seem to report the highest multipliers, while those from RBC models and SEE approaches seem to be lowest. While the means are in a range of 0.5 to 1.1, one should be aware that they comprise all kinds of fiscal impulses. Histograms for each category are displayed in Figures 1 and 2.

Since multipliers are calculated non-uniformly, Figure 3 reports average multipliers and their standard deviations for different horizons and whether they are calculated as peak or cumulative multipliers. In our data set, peak multipliers only occur in the first six years after the fiscal shock whereas cumulative multipliers are reported on longer horizons. Peak multipliers seem to be higher on average for all horizons and they are

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

fiscal impulse					
	TOTAL	GSPEND	TRANS	TAX	DEFICIT
Mean	0.85	1.01	0.39	0.54	0.31
Median	0.79	1.00	0.25	0.36	0.34
Std. dev.	0.77	0.78	0.47	0.66	0.43
Max	3.90	3.90	2.10	3.57	1.10
Min	-1.75	-1.75	-0.83	-1.50	-0.29
DH p	0.00	0.00	0.00	0.00	0.37
N	1063	721	59	240	24
fiscal impulse					
	SPEND	CONS	INVEST	MILIT	EMPLOY
Mean	1.00	0.80	1.36	1.07	0.90
Median	1.00	0.89	1.15	0.98	0.95
Std. dev.	0.79	0.57	0.89	0.82	0.77
Max	3.90	3.00	3.80	3.56	2.93
Min	-1.70	-1.75	0.00	-0.43	-0.61
DH p	0.00	0.00	0.00	0.03	0.03
N	412	173	112	24	19
model class					
	DSGE-NK	RBC	MACRO	SEE	VAR
Mean	0.76	0.55	1.05	0.58	1.00
Median	0.69	0.49	1.00	0.45	0.95
Std. dev.	0.66	0.78	0.48	0.78	0.85
Max	3.90	2.50	2.50	3.08	3.73
Min	-0.83	-1.50	0.20	-0.75	-1.75
DH p	0.00	0.19	0.07	0.00	0.00
N	358	54	92	119	440

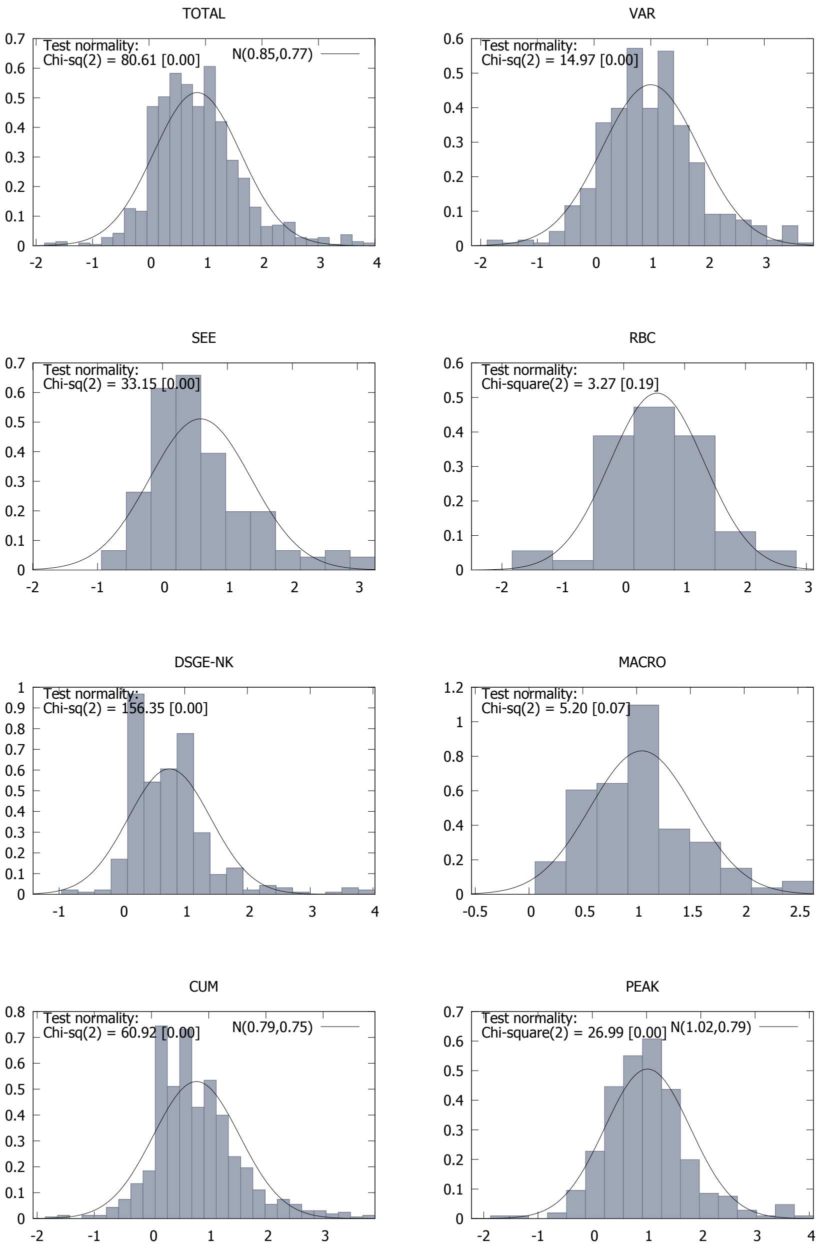

Figure 1: Histograms of reported multiplier values for total sample, various model classes and peak vs. cumulative calculation

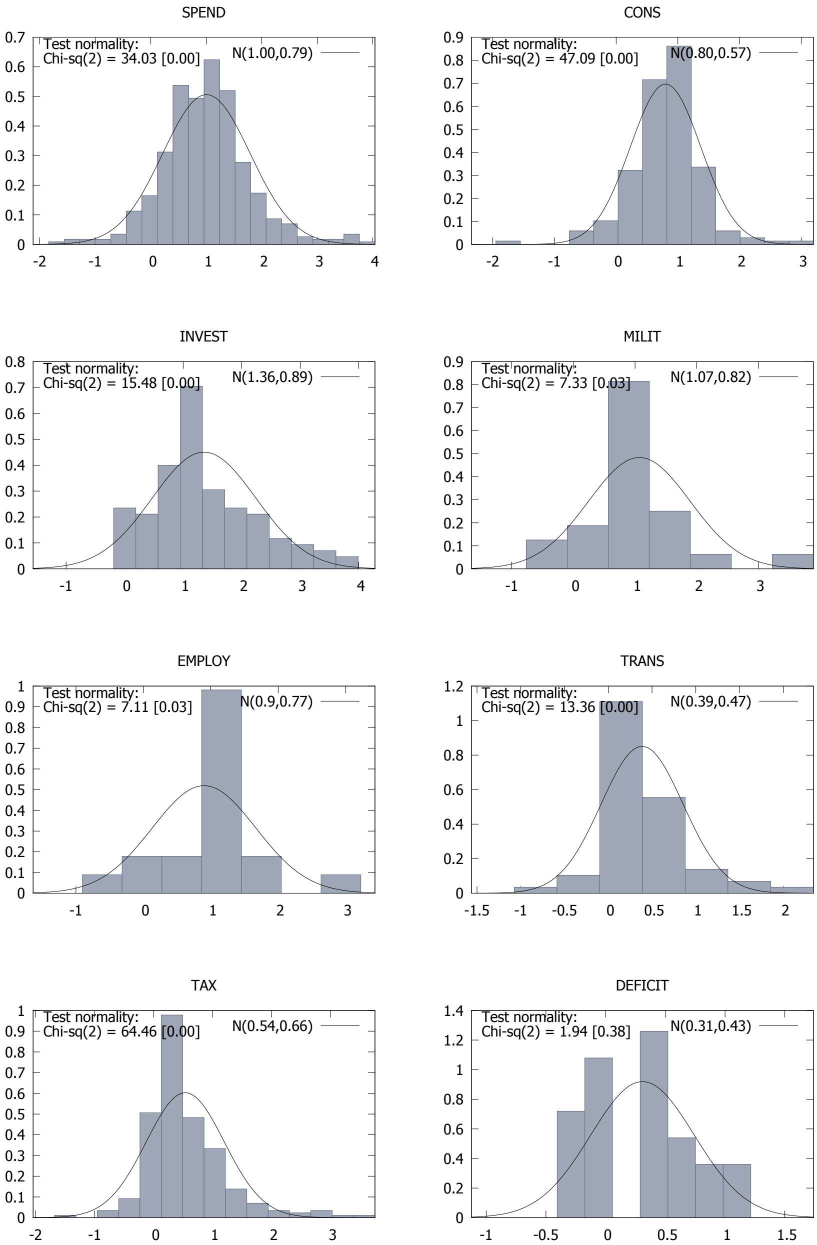

Figure 2: Histograms of reported multiplier values for various fiscal impulses

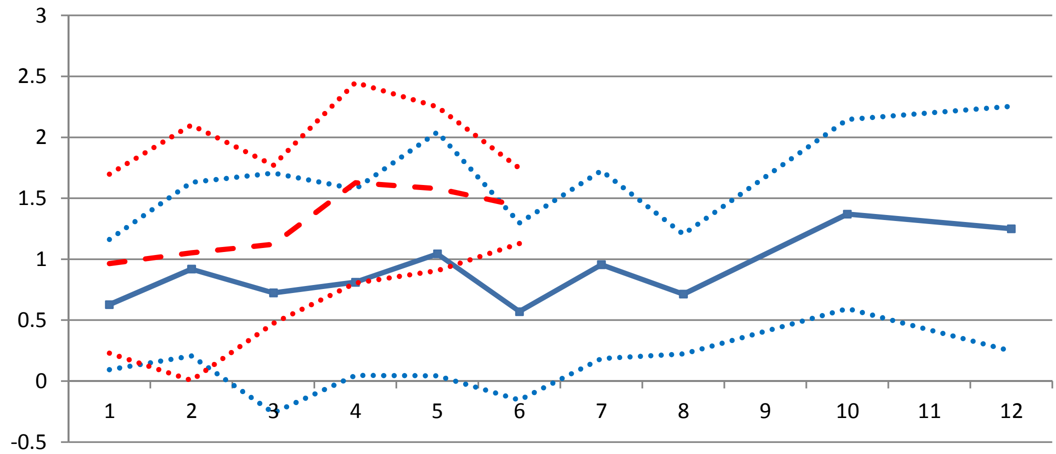

Figure 3: Mean cumulative (blue-solid) and peak (red-dashed) multipliers for different horizons in years with one-standard deviation bands (dotted)

highest after four years. Cumulative multipliers show a slightly increasing trend. Note that cumulative multipliers do not necessarily represent the shape of the IRF of GDP since they also carry information about the shape of the IRF of the fiscal impulse.

It should be stressed that we do not claim that these statistics represent true multiplier values, even if they stem from a comprehensive literature survey. Potentially significant influences that are under discussion in the literature are not included here and will not be dealt with before Section 6. Moreover, multiplier calculations may all be biased by factors that have not yet been taken into account by this literature. Properties of the distribution advise caution as well. Even though means are close to medians for each model class and impulse, Figures 1 and 2 show that multipliers of the subgroups are by and large not normally distributed, which is confirmed by Doornik-Hansen probabilities (DH p). Multimodal distributions point to additional influential factors. Obvious distortional factors are model classes interfering with fiscal impulses and *vice versa*, but also additional factors in the study design could affect the distribution. The following sections thus develop a multivariate meta regression model to check the significance and quantify the influence of fiscal impulses and study-design characteristics.

4 Meta Regression – Method

Some methodical issues need to be addressed first. Meta studies often normalize the effect size. Normalization 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 standardized manner. We control for the multiplier calculation method and the time horizon to extract comparable multiplier values. There is no other established method to translate, for example, peak multipliers into cumulative multipliers, or a multiplier for a horizon of ten quarters into a multiplier of five quarters.

According to Goldfarb and Stekler (2002), a general problem is double counting when several studies use the same data set (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 purposes, 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 using only one observation per study or taking 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 between variability and degrees of freedom. Second, when picking only one observation per study, the meta analyzer 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 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 other authors (De Grauwe and Costa Storti 2004; Nijkamp and Poot 2004; Card et al. 2010; Rusnák et al. 2011), therefore prefer to include more than one observation per study. This method has been shown to be superior to picking single observations per paper (Bijmolt and Pieters 2001). By using dummies for each paper, the specifics of a study is controlled for to a certain degree.

Nevertheless, we are aware of the problem of over-weighting, and thus check the robustness of this choice: first, by excluding single papers with many observations ($N > 30$) from our sample; second, by taking only one observation per study into account, namely the median value; third, for each (sub-)sample by setting up a weighted sample, weighting each observation of a paper by the inverse of the number of observations in the paper, thus giving each study an equal weight (Sethuraman 1995). By doing so, we strike a balance between proportional influence of single studies versus degrees of freedom and variability in our survey. The resulting coefficients are weighted least squares estimators.

A usual exercise for meta analyses is to control for a possible publication bias, i. e. the preference for statistically significant and theory-compliant results in publication selection (Stanley 2008). A standard assumption is that the results of an unbiased literature should center symmetrically around its most precise estimations. Since most of the studies we included lack comparable standard errors of their multiplier estimations, we are unable to perform standard publication bias tests via the inverse of the standard error as a measure of precision. Hence, as a second-best proxy for precision, we rely on the number of observations used for each multiplier estimation (Stanley and Doucouliagos 2012: 73). Figure 4 reports the funnel graph of our measure of precision against the reported multiplier value for the sample of empirically-based observations, that is, from estimated RBC, DSGE-NK and MACRO models and all VAR and SEE observations.

Table 2 shows several test regressions for funnel-asymmetry (Stanley and Doucouliagos 2012: 62), where the reported multiplier is regressed on various transformations of the number of observations that was used for the respective multiplier estimation ($f(obs)$). For columns (1) through (4) high significance of $f(obs)$ would point to publication selection bias. For columns (5) and (6), where the dependent variable is weighted by the log and the square root of the number of observations respectively, a publication selection bias would be indicated by a significant intercept. All of the tests do not reject the null of no asymmetry. Therefore, our sample does not point to publication selection bias at first sight, but we reevaluate this finding after controlling for additional factors in Section 6. Basically, we do not expect a systematic preference for significant positive or negative multipliers, since the different approaches in multiplier theory provide arguments for a broad band of possible results. Moreover, multipliers are usually calculated irrespective of their significance levels against zero.

The significant intercepts in columns (1) through (4) as well as the significantly positive coefficients of the measure of precision in columns (5) and (6) of Table 2 point to a genuine positive underlying multiplier effect of about 0.85 for the sample of empirically-based multipliers – in line with the mean of multipliers for the total sample.

As has been argued above, reported multiplier values are supposed to be influenced by several structural and study-design characteristics. The next section develops the regression model and a set of moderator variables that capture the study-design characteristics and structural characteristics that gave rise to discussions in the literature.

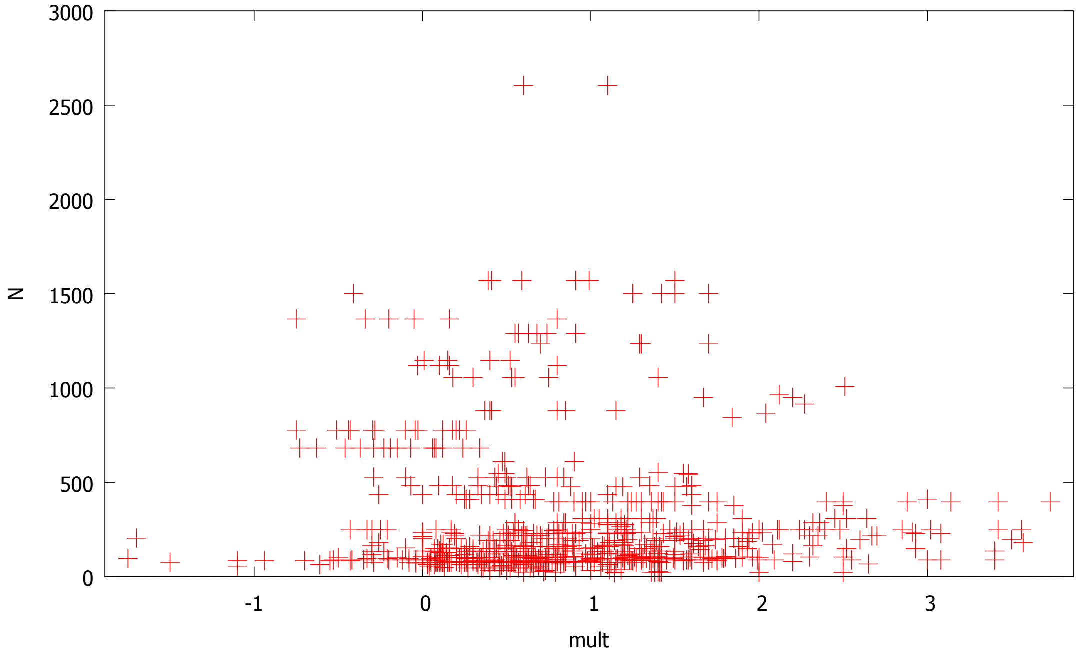

Figure 4: Funnel graph of empirical multiplier estimations

Table 2: Tests for publication selection bias

	(1)	(2)	(3)	(4)	(5)	(6)
κ	0.8555***	0.8554***	0.8553***	0.8561***	0.02443	-0.06187
	(29.53)	(29.52)	(29.39)	(29.41)	(1.008)	(-0.3067)
$f(obs)$	0.03031	-0.0009670	-0.0001273	0.0005266	0.7253***	0.6962***
	(1.001)	(-0.2880)	(-0.09001)	(0.4870)	(16.77)	(13.35)
N	746	746	746	746	746	746
$Adj.R^2$	0.0012	0.0001	0.0000	0.0001	0.4986	0.4965
ℓ	-882.4	-882.8	-882.8	-882.8	-747.5	-2328

(1) Dependent: multiplier. Indep: $f(obs) = log(obs)$.
(2) Dependent: multiplier. Indep: $f(obs) = sqrt(obs)$.
(3) Dependent: multiplier. Indep: $f(obs) = 1/log(obs)$.
(4) Dependent: multiplier. Indep: $f(obs) = 1/sqrt(obs)$.
(5) Dependent: mult$\cdot log(obs)$. Indep: $f(obs) = log(obs)$.
(6) Dependent: mult$\cdot sqrt(obs)$. Indep: $f(obs) = sqrt(obs)$.
*, **, *** indicate significance at the 10, 5, 1 percent level respectively, t-values in parentheses

5 Meta Regression – Moderator Variables

For the proposed meta regression analysis we refer to Stanley and Jarrell (2005: 302). Our empirical 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 being the multiplier value of observation j, κ the "underlying" or "reference" multiplier value, Z_j the vector of characteristics ("moderator variables") related to observation j, α the vector of systematic effects of Z_j on k_j, and δ_i the vector of paper-specific intercepts (paper dummies).

For each of the $M = 104$ studies, we include a dummy δ_i, termed study-level effect in meta regression analysis (Stanley and Doucouliagos 2012: 113), in order to control for possible cluster effects. This is equivalent to a fixed-effects model applied to meta analysis data. We also use heteroskedasticity-robust and cluster-robust standard errors, clustered by studies. To keep track of the main results we do not display the 104 paper dummies in regression tables, but we discuss their influence.

A multiplier observation in a study comes with specific characteristics, captured in the vector Z_j including the different kinds of fiscal impulses as well as study-design characteristics that should explain some of the variance in the reported multiplier values. To this end, we focused 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, but in subsamples we were able to check the influence of further study-design characteristics.

Most characteristics, such as the fiscal impulse itself, are measurable on a nominal scale only, i. e. there is no possible ranking order. We group these characteristics, since they are mutually exclusive. A reported multiplier value must exclusively belong to one value in the group 'fiscal impulse', which incorporates the values (SPEND, CONS, INVEST, MILIT, EMPLOY, TRANS, TAX, DEFICIT). SPEND applies when the paper reports the effect of a change in 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 changes in taxation (TAX). We do not distinguish between the various types of taxation. 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 more indirect types of fiscal impulses (TRANS, EMPLOY, TAX). DEFICIT applies, when the observation refers to a change in the public deficit without any information as to whether the shock is on the spending or revenue side of the budget. In line with the literature we regard multipliers as linear with respect to scale and sign of the fiscal shock.

Besides the fiscal impulse, for the total sample we focus on the influence of model classes (RBC, DSGE-NK, MACRO, VAR, SEE), which is also recorded on a nominal scale. Again, an observation must belong exclusively to one value in this group. Moreover, we include some control variables. We record the multiplier calculation method with a dummy for peak vs. cumulative calculation (PEAK, CUM). As pointed out, multiplier calculations also differ concerning the time horizon of measurement (Brückner and Tuladhar 2010: 16), so we record the log of the number of quarters after the shock (HORIZON) on which the multiplier calculation is based. We take the log of this variable in order to normalize its distribution because the majority of observations stem from shorter horizons. We also add a quadratic term of the horizon to allow for the usual hump-shaped behavior of impulse-responses.

By considering both the calculation method and the horizon, we can account for the effect of peak multipliers usually being recorded on a shorter horizon than cumulative multipliers. Thus, the pure method-specific effect is separated from the timing effect. By this combination, impact multipliers simply fall into the category of cumulative multipliers with horizon 1. To allow for different slopes, interaction terms (PEAK*HOR, PEAK*HOR^2) for the linear and quadratic variable of horizon with PEAK are included.

Another issue that should be controlled for is the leakage of fiscal impulses through the import channel as a country-specific effect (OECD 2009). Using the World Bank World Development 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.

Except for observations from purely calibrated models, we include the log of the number of observations the multiplier estimation is based upon (LOGOBS) such that, again, we can control for the influence of precision and a possible publication selection bias in the presence of our moderator variables. The log is preferred to the square root since it has more explanatory power in Table 2 as shown by t-stats.

We build two special subsamples providing the opportunity to test the robustness of our results for a smaller sample, and to consider more detailed characteristics that would not be comparable for the full sample. Subsample I comprises VAR and SEE approaches that share some mutual characteristics, such as time series properties and identification approaches. The time series properties added are a normalized value of the average year of the respective time series (AVYEAR), and a dummy for annual vs. quarterly data (ANNUAL, QUART). AVYEAR could provide information as to 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). Using annual or quarterly data may have implications for the identification of the shocks as well as for the dynamics of estimated impulse responses (Beetsma and Giuliodori 2011).

Estimation-based studies identify discretionary fiscal impulses with different strategies, which we record as a mutually exclusive group of dummies. We distinguish instrument variable estimations (IV), cyclically adjusted budget variables (CA), war episodes (WAR), the narrative record (NAR), the recursive approach (RA), the structural VAR approach of Blanchard and Perotti (BP) and the sign-restricted VAR approach (SR). IV and CA are applied in SEE models only, while RA, BP and SR refer to VAR models. NAR and WAR are applied in both. In order to capture possible differences of the NAR and WAR identification among VAR and SEE models, we include an interaction term (SEE*WAR).[2]

Subsample II, which is the complement to subsample I, comprising DSGE-NK, RBC and MACRO models, also allows for specific controls, namely, agent behavior, the modeling of the interest rate reaction, and whether the model is an open-economy model. With respect to agent behavior, we record the share of non-Ricardian agents (NONRICARD), for whom Ricardian equivalence is broken. The higher the share of non-Ricardian agents, the higher should be the reported multiplier. We assume MACRO models to have a share of non-Ricardian agents of 100 percent regarding their short-run behavior for reasons of comparability to the representative agent models in our data set.

The modeling of the interest rate can take one of four mutually exclusive values on a nominal scale, namely, an inflation-targeting central bank reaction function, including Taylor rules (INFTARG), via a loanable funds market (LOANABLE), a fixed real interest rate (FIXREAL), and a zero lower bound setting with a fixed nominal interest rate (ZLB), 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 are expected to 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

[2] A second interaction term (SEE*NAR) would be perfectly collinear with VAR, SEE, the group of identification variables and SEE*WAR.

closed-economy models to report higher multipliers.

A list of all variables can be found in Table 3. All non-dummy variables are demeaned in order to leave the intercept of the meta regressions, the reference value, unaffected by inclusion or exclusion of these variables.

6 Meta Regression – Results

6.1 Total Sample

We start by regressing reported multipliers of the total sample on characteristics as shown in Table 4. Apart from analyzing the relative effectiveness of fiscal impulses from the full set of observed multipliers, we test possible differences stemming from the choice of model class and check whether the more simulation-based approaches (RBC, DSGE-NK, MACRO) fit to the more estimation-based approaches (VAR, SEE). Results concerning the different types of fiscal stimuli bear the danger of being misleading due to the inclusion of model-based simulations. Simulations could be highly artificial, depending on parameter calibration, and could thus distort the "true picture" of multiplier values. However, there are plausible arguments to be raised against this notion. First, the surveying literature on fiscal multipliers also combines findings from simulations and estimations, so this is state of the art in the existing literature. Second, simulations stem to a large extent from estimated and well-established models that are used by institutions and think tanks. Column (2) of Table 4 focuses on observations from estimated models and purely empirical approaches and shows no considerable deviation from the total sample. Third, we include reported multipliers from sensitivity analyses of studies whenever possible, which should filter out some arbitrariness of calibrations. Fourth, even the fully estimation-based VAR and SEE approaches rely on a good deal of restrictions to reach identification and regarding the choice of variables for the empirical model. Fifth, to test the robustness of our findings we additionally analyze a purely empirical subsample (I).

Let us turn to the interpretation of the results of Table 4. Groups of variables measured on a nominal scale, such as model class or type of impulse, are necessarily multicollinear because they are mutually exclusive. That is why one variable from such groups is always omitted. The influence of these omitted variables is reflected in the constant (κ), which is called *reference value*. Thus, κ should not be interpreted as the *true multiplier* since it depends on the reference specification. We have tried to identify best practice specifications and take them as a reference, yet, this choice is still subjective. Coefficients of the moderator variables show deviations from the reference value, which allows us to

Table 3: List of variables for meta regression

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
TAX	tax reliefs to private sector	dummy
TRANS	transfers to private sector	dummy
EMPLOY	direct public employment	dummy
DEFICIT	unspecified tax relief or spending increase	dummy
model class		
RBC	real business cycle / New Classical D(S)GE model	dummy
DSGE-NK	New Keynesian DSGE model	dummy
MACRO	structural macroeconometric model	dummy
VAR	vector autoregressive model	dummy
SEE	single equation estimation approach	dummy
multiplier calculation method		
PEAK	calculated as peak multiplier	dummy
CUM	calculated as cumulative multiplier	dummy
HORIZON	log of horizon of the multiplier calculation	log of quarters
PEAK*HOR	interaction term PEAK and HORIZON	log of quarters
open economy leakage		
M/GDP	import quota of the surveyed country sample	percentage points
OPEN / CLOSED	open vs. closed economy model	dummy
properties of time series		
LOGOBS	log of number of obs. used	continuous
AVYEAR	average year of the series	years
ANNUAL / QUART	annual vs. quarterly data	dummy
identification strategy		
IV	instrument variable approach	dummy
CA	prior cyclical adjustment of public budget	dummy
WAR	war episodes approach	dummy
NAR	narrative record / action-based approach	dummy
RA	recursive VAR approach	dummy
BP	structural (Blanchard-Perotti) VAR approach	dummy
SR	sign restrictions VAR approach	dummy
SEE*WAR	interaction term SEE and WAR	dummy
households' behavior		
NONRICARD	share of non-Ricardian agents	percentage points
modeling of interest rate reaction		
INFTARG	loanable funds market	dummy
LOANABLE	loanable funds market	dummy
FIXREAL	fixed real interest rate	dummy
ZLB	zero lower bound / fixed nominal interest rate	dummy

make unconditional relative statements about the effectiveness of fiscal policy in a given setting as compared to an alternative setting.

Since we use a dummy for each paper to control for paper-specific intercepts, 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 run two stages of each regression. In the first step, we include all paper dummies, let the econometric software randomly choose the paper dummy to drop and calculate the mean coefficient of paper dummies. In the second step, we deliberately drop the dummy closest to this mean and therefore get a reference value with a minimized bias from paper dummies. Note that only second step regressions are shown and that the choice of the omitted paper dummy in no way influences any of the other coefficients, but only shifts the reference value κ.

The reference for the baseline estimation in column (1) is an average multiplier value calculated as a cumulative response to an unspecified public spending impulse, stemming from a VAR model. Such an observation on average reports a multiplier of 0.73 when controlling for other influences, which is significantly different from zero.

We interpret the coefficients in our regressions in the following way: Our endogenous variable, the reported multiplier, is a derivative. Coefficients of any dummy variable thus show the estimated difference of the multiplier value when the dummy is switched on, compared to the reference specification. Coefficients of continuous variables, such as horizon and the import-to-GDP ratio, show derivatives of the multiplier with respect to these independent variables.

Fiscal impulses differ significantly concerning their influence on the multiplier. Public investment yields multipliers which are significantly higher by 0.6 units as compared to the reference specification, while tax cuts and transfers have a significantly lower multiplier, about 0.3 to 0.4 units below those of direct public spending. For military spending, public employment and unspecified deficit spending there are deviations to unspecified public spending, but significance is generally low. Military spending multipliers and deficit spending multipliers are lower, while public employment has somewhat higher multipliers.

The next rows show the influences of different model classes, as compared to our baseline VAR model. RBC models report significantly lower multipliers with a difference of about 0.6, while multipliers from macroeconometric models are about 0.4 points higher, a difference which is weakly significant. Multipliers reported from New Keynesian DSGE simulations and SEE models are not significantly different in the prime specification.

Attention should also be given to control variables. Peak multipliers are, as expected,

Table 4: Total sample (Dep. Var.: multiplier)

	(1) base[a]	(2) empir[a]	(3) plain[b]	(4) rbc-ref[c]	(5) gspend-ref[d]	(6) cum[b]
κ	0.7274***	0.8001***	0.9852***	0.1426	0.8047***	0.9510***
	(3.089)	(5.155)	(3.467)	(0.5602)	(2.888)	(2.862)
fiscal impulse						
CONS	0.1122	0.08394	0.1212	0.1122		0.1320
	(0.9746)	(0.6271)	(1.033)	(0.9746)		(1.023)
INVEST	0.6169***	0.6607***	0.5964***	0.6169***		0.5889***
	(3.566)	(2.949)	(3.336)	(3.566)		(3.052)
MILIT	-0.2297	-0.1920	-0.2703*	-0.2297		-0.7185***
	(-1.418)	(-0.8540)	(-1.760)	(-1.418)		(-4.590)
TAX	-0.3208***	-0.2658*	-0.3261***	-0.3208***	-0.4385***	-0.2907**
	(-3.200)	(-1.877)	(-3.255)	(-3.200)	(-4.222)	(-2.350)
TRANS	-0.3852***	-0.3015**	-0.3832***	-0.3852***	-0.6190***	-0.4138***
	(-3.552)	(-2.220)	(-3.523)	(-3.552)	(-5.230)	(-3.315)
EMPLOY	0.2348	0.3020	0.2244	0.2348	0.06782	0.3362**
	(1.429)	(1.499)	(1.339)	(1.429)	(0.5454)	(2.072)
DEFICIT	-0.1374	-0.06360	-0.1190	-0.1374	-0.2369**	-0.09328
	(-1.171)	(-0.5292)	(-1.051)	(-1.171)	(-2.339)	(-0.7362)
model class						
VAR				0.5848**		
				(2.277)		
RBC	-0.5848**	-0.9300***	-0.5940**		-0.5466**	-1.092***
	(-2.277)	(-7.208)	(-2.296)		(-2.212)	(-2.609)
DSGE-NK	0.1888	-0.5208***	0.1358	0.7736***	0.2398	-0.08585
	(0.9513)	(-12.34)	(0.8074)	(3.559)	(1.260)	(-0.3167)
MACRO	0.4287*	0.4749***	0.4735**	1.013***	0.4722**	0.1195
	(1.762)	(3.884)	(2.302)	(3.905)	(2.003)	(0.3704)
SEE	0.08541	0.2987*	-0.2309	0.6702***	0.07776	0.04169
	(0.4308)	(1.799)	(-1.303)	(2.989)	(0.4072)	(0.1174)
control variables						
PEAK	0.3362**	0.3168**		0.3362**	0.3247**	
	(2.477)	(2.415)		(2.477)	(2.336)	
HORIZON	0.1670***	0.1754***		0.1670***	0.1567***	0.1810***
	(4.623)	(4.577)		(4.623)	(4.413)	(4.779)
HORIZON^2	-0.002467	-0.005844		-0.002467	-0.005850	-0.02333
	(-0.06117)	(-0.1295)		(-0.06117)	(-0.1441)	(-0.5131)
PEAK*HOR	0.01155	0.02035		0.01155	0.01866	
	(0.1229)	(0.2108)		(0.1229)	(0.2105)	
PEAK*HOR^2	0.08584	0.09038		0.08584	0.07799	
	(1.167)	(1.110)		(1.167)	(1.084)	
M/GDP	-0.01033***	-0.01436***		-0.01033***	-0.01056***	-0.009834***
	(-3.301)	(-2.938)		(-3.301)	(-3.380)	(-2.812)
LOGOBS		0.2402				
		(0.9323)				
N	1063	746	1063	1063	1063	766
Adj.R^2	0.4138	0.4022	0.3766	0.4138	0.3750	0.4278
ℓ	-882.7	-640.7	-918.8	-882.7	-918.5	-604.1

[a] reference: SPEND, VAR, CUM
[b] reference: SPEND, VAR
[c] reference: SPEND, RBC, CUM
[d] reference: GSPEND, VAR, CUM
*, **, *** indicate significance at the 10, 5, 1 percent level, t-values in parentheses

significantly higher than cumulative multipliers by a magnitude of 0.3. Doubling the horizon of multiplier calculation significantly increases the reported multiplier by 0.12 units for the linear term.[3] The quadratic term has a plausible sign, but is insignificant. The insignificant interaction terms signal that the slopes of peak and cumulative multipliers over the range of horizons should not differ much. Import quotas have a highly significant impact on reported multipliers. A one percentage point increase in the import quota should reduce the multiplier by 0.01 units.

To check robustness, we estimated some variants of the regression of column (1). In column (2), only observations from estimated models and purely empirical approaches are taken into account. Results do not change much with the exception of the influence of model classes. Multipliers from New Keynesian DSGE and RBC models now deviate more strongly from VAR models and have a significantly negative coefficient. This is plausible since the bigger estimated DSGE models in our data set have not been used that extensively to analyze special scenarios such as zero lower bound situations as compared to the calibrated models of theoretical papers. With regards to the measure of precision (LOGOBS), we find an insignificant influence, reconfirming the result of column (1) of Table 2. However, in the presence of our moderator variables, the coefficient has increased by approximately one order of magnitude.

Column (3) 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. This is reasonable given that the influence of the higher peak multipliers is not controlled for here. This is also the reason why the SEE coefficient turns negative as there are, by definition, no PEAK multipliers from SEE approaches. Thus, the difference of peak and cumulative multipliers is picked up by SEE under this plain specification.

The regression model in column (4) tests the impact of exchanging the reference model class as compared to column (1). Using observations from RBC models as a reference only affects the intercept κ and the coefficients of the model class group. The test reveals that RBC models report multipliers which are insignificantly different from zero and significantly lower compared to any other model class.

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). When we take the mix of direct public spending as reference, the difference to tax and transfer multipliers increases, which is plausible since the

[3] Since the horizon is taken in logs, the effect of doubling is $0.167 \cdot (\log 2)$.

reference value now includes the higher multipliers from public investment.

Column (6) focuses on cumulative multipliers. Results are comparable to the full sample with some exceptions. The coefficient of military spending is rather sensitive to this subsampling with relatively low cumulative multipliers, which could be interpreted in that military spending hikes are persistent, and have a rather short-lived though intense impact on GDP. RBC models report even more negative cumulative multipliers as compared to VAR results. MACRO models are rather close to the VAR benchmark for cumulative multipliers.

Some results for the total sample are very robust across all specifications and are worth being summarized: There is a significant positive public spending multiplier of approximately 0.85 for the reference specification. Public investment multipliers are significantly higher by a magnitude of 0.6, while tax and transfer multipliers are lower in the range of 0.3 to 0.4. Other fiscal impulses are by and large not significantly different from unspecified public spending. Model classes matter with RBC models producing lower multipliers and MACRO models producing higher multipliers than the benchmark. Peak multipliers are higher than cumulative multipliers by about 0.3 and longer horizons lead to increased multipliers, a result which is in line with the habit persistence hypothesis of private demand (Boldrin et al. 2001; Brown 1952). A more intense import leakage leads to lower multipliers of 0.01 to 0.015 units per one percentage point increase of the import quota.

6.2 Estimation-based Subsample

We now turn to a purely estimation-based subsample, which takes into account observations from VAR and SEE approaches only. Results of this subsample I are shown in Table 5.

Column (1) has the same model specification as column (2) of Table 4 and results are largely equivalent. The reference multiplier, which is a cumulative multiplier stemming from a public spending impulse in a VAR model, is positive and significant. Public investment has the highest multipliers, 0.8 units above the reference specification, and indirect fiscal stimuli, such as tax reliefs and transfers, imply roughly 0.25 units lower multipliers, but are not significant at the 10 percent level. Multipliers from public employment are again considerably higher, but this time significantly so. Military spending and unspecific deficit spending imply somewhat lower multipliers. Multipliers from SEE are higher, peak multipliers exceed cumulative ones and increasing the horizon of measurement also increases multipliers with the quadratic term being implausibly positive

Table 5: Subsample I (Dep. Var.: multiplier)

	(1) base[a]	(2) add[b]	(3) plain[c]	(4) var-only[b]	(5) cum[b]
κ	0.6339***	0.9447***	0.8295***	1.108***	1.309***
	(3.110)	(5.231)	(5.416)	(3.931)	(4.891)
fiscal impulse					
CONS	0.03220	0.08131	0.1108	0.2138	0.08687
	(0.1424)	(0.3791)	(0.5041)	(0.8486)	(0.3542)
INVEST	0.7910***	0.8340***	0.8163***	1.013***	0.7957***
	(3.057)	(3.555)	(3.362)	(4.652)	(3.013)
MILIT	-0.1455	-0.3464	-0.2656	-0.03200	-1.118**
	(-0.6135)	(-0.8933)	(-0.8913)	(-0.2987)	(-3.195)
TAX	-0.2431	-0.2647	-0.2578	-0.4284***	-0.1564
	(-1.375)	(-1.482)	(-1.476)	(-2.659)	(-0.6899)
TRANS	-0.2240	-0.2638	-0.1328	-0.2579**	-0.2028
	(-1.288)	(-1.454)	(-0.7204)	(-2.419)	(-0.6042)
EMPLOY	0.4317**	0.6369***	0.4614**	0.6824***	0.7197***
	(1.982)	(3.274)	(2.059)	(2.648)	(3.405)
DEFICIT	-0.04852	-0.02710	-0.005811	-0.5514*	0.008594
	(-0.3848)	(-0.2767)	(-0.05781)	(-1.740)	(0.07800)
model class					
SEE	0.3145*	0.6398***	0.3416***		0.6048***
	(1.767)	(5.249)	(3.053)		(24.22)
control variables					
PEAK	0.3770***	0.3778***		0.3655***	
	(2.929)	(2.786)		(2.644)	
HORIZON	0.1814***	0.1776***		0.1594***	0.1696***
	(4.573)	(4.387)		(3.532)	(3.628)
HORIZON^2	0.001272	-0.004834		-0.01142	-0.07239
	(0.02517)	(-0.09240)		(-0.2070)	(-1.194)
PEAK*HOR	0.08417	0.07563		0.1011	
	(0.9247)	(0.8412)		(1.097)	
PEAK*HOR^2	0.04388	0.02142		0.01843	
	(0.5265)	(0.2453)		(0.2092)	
M/GDP	-0.02252***	-0.02005***		-0.01990***	-0.02229**
	(-2.859)	(-2.775)		(-2.845)	(-1.979)
LOGOBS	0.2597	0.3572**	0.5451***	0.4673***	0.3459
	(0.9923)	(2.154)	(2.963)	(2.934)	(1.541)
additional characteristics					
BP		-0.09155	0.1960	-0.1286	0.004109
		(-1.152)	(1.612)	(-1.440)	(0.03158)
RA		-0.1302	0.1620	-0.2129*	-0.05883
		(-1.107)	(1.216)	(-1.731)	(-0.3391)
SR		-0.5024***	-0.1824	-0.5635***	-0.3755**
		(-5.440)	(-1.290)	(-5.884)	(-2.470)
WAR		-0.5790***	-0.3480**	-0.7078***	-0.7097***
		(-3.212)	(-2.112)	(-3.924)	(-3.443)
IV		-0.2890*	-0.4690		-0.2176
		(-1.936)	(-0.9546)		(-1.150)
CA		-0.9856***	-1.039***		-0.9714***
		(-9.747)	(-9.572)		(-9.224)
SEE*WAR		0.7986**	0.5674**		1.676***
		(2.138)	(2.027)		(5.128)
ANNUAL		-0.5768***		-0.6293***	-1.469***
		(-3.173)		(-3.829)	(-3.672)
AVYEAR		-0.02511***		-0.02773***	-0.02547***
		(-4.248)		(-4.911)	(-4.060)
N	559	556	559	438	402
$Adj.R^2$	0.3974	0.4238	0.3553	0.3650	0.4366
ℓ	-526.6	-505.4	-544.9	-419.4	-360.8

[a] reference: SPEND, VAR, CUM
[b] reference: SPEND, VAR, CUM, NAR, QUART
[c] reference: SPEND, VAR, NAR
*, **, *** indicate significance at the 10, 5, 1 percent level, t-values in parentheses

but very small and insignificant. Economies with a higher share of aggregate demand going abroad face lower multipliers. The measure of precision (LOGOBS) has a positive but insignificant value, comparable to column (2) of Table 4.

In columns (2) through (6) we add control variables specific for VAR and SEE approaches, which are the different ways of identification as well as characteristics of the time series used. The reference specification here is a VAR using the narrative approach and quarterly data. We used the narrative approach since it is applicable to both VAR and SEE approaches and is among the most influential ones in the literature. Due to a severe outlier in the AVYEAR variable, we excluded the study of Almunia et al. (2010) that analyses the Great Depression years.

Column (2), which has the most comprehensive specification, generally confirms the results of column (1) – coefficients are altered only slightly with a few exceptions. The reference value is higher because of additional dummy variables regarding identification and time series properties as discussed below. This time, the measure of precision is significant, pointing to a negative publication bias with increased multipliers for more precise studies, which would translate into an increase in the multiplier of 0.25 units when doubling the sample size.

With respect to identification approaches, their results are rather distinct from each other. The reference value has increased as compared to column (1) since NAR is on the upper end of the scale. There are three clusters: first, the BP, RA, NAR and IV approaches, which represent the majority in our data set, producing the highest multipliers with coefficients being close to each other; second, the SR and WAR coefficients with multipliers being significantly lower by roughly 0.55 units; third, the method of using cyclically-adjusted budget variables which stands out from the rest with 1.0 units lower multipliers. If WAR were to be used as the reference specification, CA would still be 0.4 units lower with the coefficient being highly significant (results not shown). The interaction term SEE*WAR is also significant and signals, in sum with WAR, that for SEE models multipliers from the war episodes approach are not lower compared to the narrative approach.

Using annual data considerably decreases measured multipliers. The quarterly vs. annual dummy may contain information regarding anticipation and institutional issues of identification as discussed for example in Beetsma and Giuliodori (2011: F11). Multipliers from more recent time series are lower as well. The latter result is robust when using the first or the last year of the times series instead of the average year (results not shown). This is in line with findings in the literature of a 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.

Column (3) excludes the group of control variables and time series properties to show robustness in this regard. There are only some minor quantitative changes to column (2) with the reference value being lower since we do not control for annual vs. quarterly data here, and the impact of the publication bias being stronger. Column (4) takes only VAR models into account. As compared to column (2), the reference multiplier increases a little. Turning to fiscal impulses, investment multipliers are even more distinct in VAR models and multipliers from taxes, transfers and unspecific public deficits are now significantly lower than those of public spending impulses. The recursive approach yields slightly lower multipliers, this time significant. The coefficient of WAR is bigger for the group of VAR models, in line with what has been said above. Other coefficients are very close to those in column (2).

In column (5) we turn to a sample of cumulative multipliers only, which yields some differences. Cumulative multipliers of a spending shock for the narrative VAR reference specification are higher than in column (2). Military spending, as in column (6) of Table 4, has low cumulative multipliers. The difference of tax and transfer multipliers to those of public spending is less pronounced for cumulative multipliers. The coefficient of LOGOBS turns insignificant but does not change much in terms of quantity. Identification with the WAR dummy approach has lower cumulative multipliers in VAR models, but much higher multipliers for SEE models. The influence of annual as opposed to quarterly data is more pronounced for cumulative multipliers, while the coefficient of AVYEAR roughly stays the same.

We may summarize the most robust findings across all specifications of subsample I. There is a significant positive public spending multiplier for the reference specification, though this depends on the reference identification method. If we refer to the methods which make up the lion's share of multiplier estimations in the literature as best practice, then the reference multiplier is close to 1 for our subsample. Public investment multipliers are considerably higher by a magnitude of 0.9, those from public employment are higher by roughly 0.6. Tax and transfer multipliers are lower by about 0.25 units, but with varying significance. Military spending multipliers are lower by about 0.2 units, much more pronounced for cumulative multipliers.

SEE models report higher multipliers than VAR models. Peak multipliers are higher than cumulative multipliers by about 0.35 units and longer horizons lead to increased multipliers, a result which is in line with the habit persistence hypothesis of private demand (Boldrin et al. 2001; Brown 1952). The high significance levels of multiplier calculation method and horizon that already appeared in Table 4 seem to have their roots in VAR and SEE approaches, since they do not appear in subsample II as will be shown below. This is perfectly in line with intuition because the IRFs of the simulation-based approaches are much smoother. Increasing the import quota by one percentage point lowers the multiplier by roughly 0.02 units. There are some signs that a publication selection bias is present in subsample I, with multipliers increasing considerably with an increased sample size. Note, however, that the number of observations is only a second-best proxy for precision.

Regarding identification, the narrative approach, the Blanchard-Perotti method, the recursive and IV approaches roughly point to the same multipliers. Multipliers from sign restriction approaches are on average 0.35 units lower, those of the war episodes approach are rather ambiguous as they differ strongly for VAR and SEE models. Using cyclically-adjusted budget variables clearly stands out with very much lower multipliers of about 1.0 units. Annual data imply lower multipliers while multipliers generally seem to have declined over the past decades, a result which we are not able to test for the recent crisis years since most of the studies in our sample do not use such recent data. However, we set up a crisis specification for the simulation-based approaches in the following subsection.

6.3 Simulation-based Subsample

Subsample II is the complement to subsample I as it comprises all observations from simulation-based approaches (RBC, DSGE-NK, MACRO). Regression results are shown in Table 6, which provides the prime regression for this subsample as well as some robustness checks with additional variables applicable for these model classes.

Most results of the total sample are reaffirmed with subsample II. The reference specification in column (1) stems from a cumulative multiplier estimation of a public spending shock in a DSGE-NK model. Such a specification produces a significantly positive multiplier value of about 1.0. Investment multipliers are still higher than those of public spending, even though the difference is not significant. In model simulations, military spending yields much higher multipliers than estimated for the full sample and subsample I. Tax cuts and transfers have multipliers that are significantly lower by 0.55 units. Other fiscal stimuli are close to the reference specification.

Concerning model classes there is a significant difference among them in the prime specifications with RBC models yielding multipliers close to zero and MACRO models having multipliers above one. The control variables for the shape of impulse-responses

Table 6: subsample II (Dep. Var.: multiplier)

	(1) prime[a]	(2) add[b]	(3) dsge-nk[b]	(4) cum[b]	(5) crisis[c]	(6) empir[c]
κ	0.9759***	0.8109***	0.6784***	0.9084***	1.869***	2.066***
	(6.210)	(3.377)	(10.23)	(4.886)	(7.366)	(3.458)
fiscal impulse						
CONS	0.07335	0.04738	0.2131	0.01397	0.04738	-0.2143
	(0.7397)	(0.4213)	(1.361)	(0.1369)	(0.4213)	(-1.544)
INVEST	0.2499	0.2553*	0.5017**	0.2372	0.2553*	-0.2241
	(1.621)	(1.654)	(2.380)	(1.471)	(1.654)	(-1.236)
MILIT	0.7088***	0.8093***		0.9665**	0.8093***	
	(4.723)	(3.123)		(2.434)	(3.123)	
TAX	-0.4711***	-0.4867***	-0.3801***	-0.5238***	-0.4867***	-0.6665***
	(-5.256)	(-5.607)	(-3.404)	(-5.626)	(-5.607)	(-5.102)
TRANS	-0.5966***	-0.5905***	-0.4702***	-0.6551***	-0.5905***	-0.7272***
	(-6.078)	(-6.036)	(-4.071)	(-6.262)	(-6.036)	(-5.400)
EMPLOY	-0.06832	-0.008740	0.1594	0.07156	-0.008740	-0.1876
	(-0.2795)	(-0.03227)	(0.4860)	(0.3004)	(-0.03227)	(-0.6453)
model class						
RBC	-0.8833***	-0.7315***		-1.286***	-0.7315***	-0.9966**
	(-12.21)	(-5.012)		(-27.38)	(-5.012)	(-2.186)
MACRO	0.3250***	-0.003328		0.2023	-0.003328	0.06433
	(3.097)	(-0.01010)		(0.9197)	(-0.01010)	(0.1225)
control variables						
PEAK	-0.03461	-0.03649	-0.09948		-0.03649	-0.1377
	(-0.09100)	(-0.1019)	(-0.2318)		(-0.1019)	(-0.4619)
HORIZON	0.04063	-0.01906	-0.03171	0.07354	-0.01906	-0.03964
	(0.5529)	(-0.3269)	(-0.4010)	(1.265)	(-0.3269)	(-0.1640)
HORIZON^2	0.03686	0.06143	-0.03125	0.07577*	0.06143	0.04998
	(0.7389)	(1.365)	(-0.2087)	(1.805)	(1.365)	(0.3191)
PEAK*HOR	-0.1691	-0.01016	0.03987		-0.01016	-0.6670*
	(-0.4093)	(-0.03065)	(0.1287)		(-0.03065)	(-1.857)
PEAK*HOR^2	0.02885	0.05579	0.1572		0.05579	-0.2099
	(0.09822)	(0.2227)	(0.5392)		(0.2227)	(-0.9948)
M/GDP	-0.006306***	-0.008669***	-0.007738***	-0.007875***	-0.008669***	-0.009093***
	(-3.138)	(-6.072)	(-6.824)	(-7.157)	(-6.072)	(-15.36)
additional characteristics						
NONRICARD		0.005141	0.005393	0.002083	0.005141	0.006672
		(1.267)	(1.323)	(0.7319)	(1.267)	(0.8282)
LOANABLE		-0.3165**	-0.7204	-0.3494**	-0.8275***	
		(-2.382)	(-1.122)	(-2.240)	(-4.054)	
FIXREAL		0.09681		0.07783	-0.4142**	-0.9924**
		(1.445)		(1.049)	(-2.243)	(-1.986)
ZLB		0.5110**	0.5393**	0.3623***		
		(2.553)	(2.374)	(2.608)		
CLOSED		0.3925***	0.4136***	0.3919***		
		(3.607)	(4.313)	(4.000)		
INFTARG					-0.5110**	-0.3477*
					(-2.553)	(-1.740)
OPEN					-0.3925***	0.1366
					(-3.607)	(0.3478)
LOGOBS						0.8846
						(1.272)
N	504	504	358	362	504	187
Adj.R^2	0.5016	0.5838	0.5885	0.6140	0.5838	0.7108
ℓ	-288.9	-240.6	-171.8	-121.8	-240.6	-5.317

[a] reference: SPEND, DSGE-NK, CUM

[b] reference: SPEND, DSGE-NK, CUM, INFTARG, OPEN

[c] reference: SPEND, DSGE-NK, CUM, ZLB, CLOSED

*, **, *** indicate significance at the 10, 5, 1 percent level, t-values in parentheses

are all insignificant and have negligible values. This result withstands a stepwise exclusion of the interaction terms and the quadratic term (results not shown). As mentioned above, this is reasonable since IRFs of the fiscal shock and the GDP response in model simulations die out in a similar manner and are generally smooth. From our reading of the literature, it is not state of the art to tackle the question of different timing of impulse-responses – an issue that is important for interpreting the results from VAR and SEE approaches. Import quotas, again, have a plausible negative significant impact on the multiplier.

Column (2) adds three model specific characteristics, which are the share of Ricardian vs. non-Ricardian agents, the group of interest rate reaction functions, and whether the model simulates a closed or an open economy. An increase in the share of non-Ricardian agents of one percentage point plausibly raises the reported multiplier by half a percentage point, but the coefficient is not statistically significant.

The various mutually exclusive interest rate reaction functions matter as well. The reference model has an inflation-targeting or Taylor-rule setting. Loanable funds models report significantly lower multipliers (-0.3) while those with a zero lower bound setting or a fixed nominal interest rate report multipliers that are higher by 0.5 units. Fixed real interest rates do not change multipliers much. Closed economy models plausibly increase multipliers by 0.4 units.

With the additional variables come some changes in the baseline moderator variables. The reference value is lowered somewhat, because it now distinguishes open from closed economy models and stems from an inflation-targeting model whose multipliers are a little below average. Investment multipliers now become significantly positive, however, in general, coefficients of fiscal impulses alter only slightly. An interesting, though plausible change is that a stepwise inclusion[4] of additional characteristics reveals that the difference between DGSE-NK and MACRO models vanishes when controlling for agent behavior, which we have defined as fully non-Ricardian for MACRO models.

Results from a plain model without the insignificant standard controls for PEAK, HORIZON and M/GDP are not reported here due to space constraints. There are only negligible differences to column (2). Coefficients are very robust against a general-to-specific exclusion of insignificant variables. In column (3) we focus on DSGE models and find most of our earlier results confirmed, albeit the reference fiscal impulse of an unspecified public spending shock seems to be about 0.15 units lower as compared to column (2), which is reflected both in the lower reference value and the altered differences to the other fiscal impulses, resulting for example in a highly significant difference for

[4]Single steps are not reported.

investment multipliers, and a lower, yet still significant difference for tax and transfer multipliers. Moreover, probably due to a small sample issue, the loanable funds setting becomes more negative but insignificant.

When looking at cumulative multipliers only (column (4)), there are only some slight quantitative changes, with RBC models reporting even lower multipliers and the zero lower bound effect and the influence of agent behavior decreasing somewhat. The quadratic term of HORIZON becomes significantly positive which would point to increasing multipliers in a longer-run perspective, a rather implausible result for model simulations.

Column (5) deals with the question of multipliers in times of crisis. It reconstructs a crisis scenario yielding an average crisis multiplier from the whole set of models in our data base. The reference specification is a zero lower bound setting with a share of non-Ricardian agents being 30 percent above average, a reasonable amount discussed in the literature (Roeger and Veld 2010); moreover, we assume a closed economy model, which should simulate a multi-country crisis fostering concerted actions of governments as observed at the beginning of the financial crisis (IMF 2010). The crisis setting could as well apply to the concerted austerity actions recently faced by the Euro Area. Such a reference multiplier as derived from a broad set of model simulations is close to 2, pointing to a strong impact of fiscal policy in crisis years. Column (6) doing the same exercise on the basis of all estimated models in our data set confirms this finding.

Again we may draw some more general conclusions from robust findings across Table 6. The reference multiplier is about 0.85 for a non-crisis specification of models and more than doubles when we depict average results from model simulations of a crisis situation. Regarding the relative effectiveness of fiscal impulses there is a less pronounced role for public investment as compared to Table 5, however, the multiplier is still about 0.2 units higher for such an impulse. Model simulations give much more credit to military spending shocks, which is not confirmed by the estimation-based approaches. Tax and transfer multipliers in subsample II are significantly lower than those from unspecific public spending by 0.5 to 0.6 units – a result which is less pronounced but roughly confirmed by subsample I.

RBC models clearly report lower multipliers close to zero and there is no difference between DSGE-NK and MACRO models when controlling for agent behavior. The multiplier calculation method and the horizon of measurement do not seem to play a role in model simulations. These issues are, however, important for estimation-based approaches which raises questions as to the ability of simulations to predict the timing of the effects of fiscal policy.

Plausibly, higher import quotas, in line with subsample I, decrease the multiplier and closed economy models predict higher multipliers. Increasing the share of non-Ricardian agents by one percentage point insignificantly raises the reported multiplier by approximately half a percentage point. When the interest rate reaction function is determined by a loanable funds market, multipliers are 0.3 units lower than for an inflation-targeting setting. A zero lower bound situation or a fixed nominal interest rate leads to higher multipliers, increased by approximately 0.5 units.

6.4 Further Robustness Tests

Before we conclude, we would like to refer to the statistical appendix that contains further robustness tests concerning a possible overweighting of comprehensive studies. The columns of Table 8 show results for the prime specification of our total sample when dropping single papers with many observations ($N > 30$) from the sample. There are hardly any changes to the results of the total sample.

In Table 9 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. Note that interpreting the magnitude of coefficients is not straightforward in this case. Generally, significance levels are lower while adjusted R^2 values are large. This is due to the fact that the paper-specific intercepts now carry the bulk of information as each paper has its specific weight. Nevertheless, signs of the coefficients are equal to the unweighted counterparts.

Finally, Table 10 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. Most results are confirmed, but there are deviations regarding public investment multipliers and peak multipliers. It should, however, be pointed out that this strategy eats away a lot of information and certainly biases the coefficients of characteristics that are usually run against each other in one study, such as the fiscal impulse or the multiplier calculation method. An observation can only become the median if it is not among the higher or lower values in a study. Given that, for example, public investment multipliers are usually among the largest in a paper, they do not become the median observation unless they are of average effectiveness. The same applies to peak multipliers.

In column (2) and (3) a probit and logit model are tested for the full sample. The dependent variable is binary, signaling whether the multiplier is greater than or equal to one or whether it is less than one. The value of coefficients are not interpretable as deviations, however, while signs and significance levels are similar to baseline estimations. The results of these robustness checks largely affirm our prime specifications.

7 Conclusions

We tested a set of 104 studies on multiplier effects by meta regression analysis in order to provide a systematic overview of the different approaches, in order to derive stylized facts on influential factors and to quantify the differing effectiveness of the composition of fiscal impulses, adjusted for the interference of study-design characteristics and sample specifics. The method is not suitable for distilling an absolute true value of the multiplier, but it is able to extract relative differences between fiscal impulses and study characteristics.

The following broad picture can be drawn from the meta analysis: First, multipliers from public spending are significantly positive and on average close to one, yet, they vary a lot with study design and the underlying sample.

Second, direct public demand impulses tend to have higher multipliers than tax cuts and transfers, even though the difference is not significant in all instances. Public investment seems to be the most effective fiscal impulse, a result that is robust against many specifications. Military spending is preferred solely by the simulation-based approaches. 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. Public employment has somewhat higher multipliers than public spending in general with a high significance among estimation-based approaches only.

Third, reported multipliers depend on model classes. Controlling for additional variables reveals that RBC models come up with significantly lower multipliers not far above zero, especially when focusing on cumulative multipliers. Backward-looking macroeconometric models report significantly higher multipliers than the reference VAR models. Multipliers from single equation estimations also seem to have somewhat higher multipliers when controlling for additional factors. DSGE-NK models report multipliers fairly close to those of VAR models, though this finding does not apply to the bigger estimated DSGE-NK models whose multipliers are significantly lower since they have not been used that extensively to analyze special scenarios such as zero lower bound situations as compared to the calibrated models of theoretical papers.

Fourth, reported multipliers strongly depend on the method and horizon of calculating them, especially for estimation-based approaches. Peak multipliers are on average 0.3 units greater than cumulative multipliers and the longer the horizon of measurement, the higher is the multiplier. Thus, a simple listing of multiplier values without additional information on how they were computed could provide a biased picture. Moreover, multipliers from simulation-based approaches are largely insensitive to this issue, which questions the ability of simulations to predict the timing of the effects of fiscal policy.

Fifth, the more open the import channel of an economy, the lower seems to be the multiplier, a fact which is correctly represented by simulation-based approaches when separating open-economy and closed-economy models.

Sixth, when controlling for fiscal impulses and study-design characteristics there are some weak signs that more precise studies report higher multipliers. Following Stanley (2008), this could point to a negative publication selection bias, but it should be stressed that our measure of precision is only second-best.

Seventh, the various identification strategies to deal with endogeneity of fiscal impulses in VAR and SEE models report rather different multipliers. Results from the majority of observations in the literature stemming from the narrative record, the Blanchard-Perotti method, the recursive and IV approaches roughly point to the same multipliers at the upper end of the scale, close to one. Multipliers from sign restriction approaches are on average 0.35 units lower, those of the WAR episodes approach are rather ambiguous as they differ strongly for VAR and SEE models. Using cyclically-adjusted budget variables clearly stands out with multipliers close to zero.

Eighth, time series from more recent years tend to yield lower multipliers, confirming the findings in 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 an adequate portion of the effects of the stimulus packages in response to the Great Recession.

Ninth, setting up a crisis scenario with a fixed nominal interest rate (ZLB), lowered ability of consumption smoothing (increased share of non-Ricardian agents) and a concerted action of the country under investigation and its trading partners (closed economy assumption) for the average of the whole set of models in our simulation-based sample yields multipliers close to two, which implies a higher effectiveness of fiscal policy in times of the recent crisis years and stronger negative impact of the concerted austerity measures in the Euro Area.

As an overall conclusion, reported multipliers very much depend on the setting and method chosen. Thus, economic policy consulting based on a certain multiplier study should signal, how strongly results are influenced by specification. Our meta analysis may provide guidance concerning such influential specifications and their direction and scale.

References

Acconcia, A., G. Corsetti, and S. Simonelli (2011). Mafia and Public Spending: Evidence on the Fiscal Multiplier from a Quasi-experiment. CSEF working paper 281.

Afonso, A., H. P. Grüner, and C. Kolerus (2010). Fiscal policy and growth: do financial crises make a difference? European Central Bank working paper series 1217.

Aiyagari, S. R. and L. J. Christiano (1992). The output, employment, and interest rate effects of government consumption. *Journal of Monetary Economics 30*, 73–86.

Alesina, A. and S. Ardagna (2010). Large changes in fiscal policy: taxes versus spending. *NBER/Tax Policy & the Economy 24*(1), 35–68.

Alesina, A. and S. Ardagna (forthcoming). The Design of Fiscal Adjustments. In J. R. Brown (Ed.), *Tax Policy and the Economy*, Volume 27. Chicago: University of Chicago Press.

Almunia, M., A. S. Bénétrix, B. Eichengreen, K. H. O'Rourke, and G. Rua (2010). From Great Depression to Great Credit Crisis: Similarities, Differences and Lessons. *Economic Policy 25*(62), 219–265.

Arcangelis, G. d. and S. Lamartina (2003). Identifying fiscal shocks and policy regimes in OECD countries. European Central Bank working paper series 281.

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 4*(2), 1–27.

Barrell, R., J. Byrne, S. Gottschalk, I. Hurst, and D. van Welsum (2004). Macroeconomic policy in Europe: experiments with monetary responses and fiscal impulses. *Economic Modelling 21*, 877–931.

Barrell, R., D. Holland, and I. Hurst (2012). Fiscal multipliers and prospects for consolidation. *OECD Journal: Economic Studies 2012*(1), 71–102.

Barro, R. J. and C. J. Redlick (2011). Macroeconomic Effects from Government Purchases and Taxes. *Quarterly Journal of Economics 126*(1), 51–102.

Baum, A. and G. B. Koester (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 03/2011.

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

Bayoumi, T. and S. Sgherri (2006). Mr. Ricardo's Great Adventure: Estimating Fiscal Multipliers in a Truly Intertemporal Model. IMF Working Paper WP/06/186.

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.

Beetsma, R., M. Giuliodori, and F. Klaassen (2006). Trade spill-overs of fiscal policy in the European Union: a panel analysis. *Economic Policy 41*(48), 639–687.

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.

Bénétrix, A. S. and P. Lane (2009). Fiscal Shocks and the Real Exchange Rate. IIIS Discussion Paper 286.

Biau, O. and E. Girard (2005). Politique budgétaire et dynamique économique en France : l'approche VAR structurel. *Revue Économique 56*, 755–764.

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.

Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001). Habit Persistence, Asset Returns, and the Business Cycle. *American Economic Review 91*(1), 149–166.

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.

Brown, T. M. (1952). Habit Persistence and Lags in Consumer Behaviour. *Econometrica 20*(3), 355–371.

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.

Burriel, P., F. d. Castro, D. Garrote, E. Gordo, J. Paredes, and J. J. Perez (2010). Fiscal Policy Shocks in the Euro Area and the US: An Empirical Assessment. *Fiscal Studies 31*(2), 251–285.

Caldara, D. and C. Kamps (2006). What Do We Know About the Effects of Fiscal Policy Shocks? A Comparative Analysis. Computing in Economics and Finance 2006 257.

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.

Caldara, D. and C. Kamps (2012). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board 2012-20.

Candelon, B. and L. Lieb (2011). Fiscal Policy in Good and Bad Times. Technical report.

Canova, F. and E. Pappa (2007). Price Differentials in Monetary Unions: The Role of Fiscal Shocks. *Economic Journal 117*(520), 713–737.

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.

Castro Fernández, F. d. and P. d. Hernández Cos (2008). The economic effects of exogenous fiscal shocks in Spain: a SVAR approach. *Journal of Macroeconomics 30*, 1005–1028.

Christiano, L. J., M. Eichenbaum, and S. Rebelo (2009). When is the Government Spending Multiplier Large? NBER working paper 15394.

Christoffel, K., G. Coenen, and A. Warne (2008). The New Area-Wide Model of the euro area: a micro-founded open-economy model for forecasting and policy analysis. European Central Bank working paper series 944.

Church, K. B., J. E. Sault, S. Sgherri, and K. F. Wallis (2000). Comparative Properties of Models of the UK Economy. *National Institute Economic Review 171*, 106–122.

Cloyne, J. (2011). What are the effects of tax changes in the United Kingdom? New evidence from a narrative evaluation. CESifo working paper 3433.

Coenen, G., C. J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Lindé, A. Mourougane, D. Muir, S. Mursula, C. d. Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt, and J. i. t. Veld (2010). Effects of Fiscal Stimulus in Structural Models. IMF Working Paper 10/73.

Coenen, G. and R. Straub (2005). Does Government Spending Crowd in Private Consumption? Theory and Empirical Evidence for the Euro Area. *International Finance 8*(3), 435–470.

Cogan, J. F., T. Cwik, J. B. Taylor, and V. Wieland (2010). New Keynesian versus Old Keynesian Government Spending Multipliers. *Journal of Economic Dynamics and Control 34*(3), 281–295.

Congressional Budget Office (2010). Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output From October 2009 Through December 2009. Technical report.

Corsetti, G., A. Meier, and G. J. Müller (2009). Fiscal Stimulus with Spending Reversals. IMF Working Paper WP/09/106.

Corsetti, G., A. Meier, and G. J. Müller (2010). Cross-Border Spillovers from Fiscal Stimulus. *International Journal of Central Banking 6*(1), 5–36.

Corsetti, G., A. Meier, and G. J. Müller (2012). Fiscal Stimulus with Spending Reversals. *Review of Economics and Statistics 94*(4), 878–895.

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

Dalsgaard, T., C. André, and P. Richardson (2001). Standard Shocks in the OECD Interlink Model. OECD Economics Department Working Papers 306.

Davig, T. and E. M. Leeper (2011). Monetary–fiscal policy interactions and fiscal stimulus. *European Economic Review 55*(2), 211–227.

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

Devereux, M. B., A. C. Head, and B. J. Lapham (1996). Monopolistic Competition, Increasing Returns, and the Effects of Government Spending. *Journal of Money, Credit, and Banking 28*(2), 233–254.

Edelberg, W., M. Eichenbaum, and J. D. M. Fisher (1999). Understanding the Effects of a Shock to Government Purchases. *Review of Economic Dynamics 2*(1), 166–206.

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

Elmendorf, D. W. and D. L. Reifschneider (2002). Short-Run Effects of Fiscal Policy with Forward-Looking Financial Markets. *National Tax Journal 55*(3), 357–386.

Erceg, C. J., L. Guerrieri, and C. Gust (2006). SIGMA: A New Open Economy Model for Policy Analysis. *International Journal of Central Banking 2*(1), 1–49.

Erceg, C. J. and J. Lindé (2010). Is there a fiscal free lunch in a liquidity trap? International Finance Discussion Papers 1003.

European Commission (2003). Public finances in EMU 2003. European Economy 3/2003.

European Commission (2012). Report on Public finances in EMU. European Economy 4/2012.

Fagan, G., J. Henry, and R. Mestre (2005). An Area-Wide Model for the Euro Area. *Economic Modelling 22*(1), 39–59.

Fair, R. C. (2010). Estimated Macroeconomic Effects of the U.S. Stimulus Bill. Cowles Foundation Discussion Paper 1756.

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.

Fernandez-Villaverde, J. (2010). Fiscal Policy in a Model with Financial Frictions. *American Economic Review 100*(2), 35–40.

Fisher, J. D. M. and R. Peters (2010). Using Stock Returns to Identify Government Spending Shocks. *Economic Journal 120*(544), 414–436.

Forni, L., L. Monteforte, and L. Sessa (2009). The general equilibrium effects of fiscal policy: Estimates for the Euro area. *Journal of Public Economics 93*(3-4), 559–585.

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.

Gechert, S. and H. Will (2012). Fiscal Multipliers: A Meta-Regression Analysis. IMK working paper 97.

Giordano, R., S. Momigliano, S. Neri, and R. Perotti (2007). The effects of fiscal policy in Italy: Evidence from a VAR model. *European Journal of Political Economy 23*, 707–733.

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

Guajardo, J., D. Leigh, and A. Pescatori (2011). Expansionary Austerity: New International Evidence. IMF Working Paper WP/11/158.

Hall, R. E. (2009). By How Much Does GDP Rise if the Government Buys More Output? *Brookings Papers on Economic Activity 2009*(2), 183–249.

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.

Heilemann, U. and H. Findeis (2009). Empirical Determination of Aggregate Demand and Supply Curves: The Example of the RWI Business Cycle Model. Universität Leipzig Working Paper 83.

Henry, J., P. d. Hernández Cos, and S. Momigliano (2008). The impact of government budgets on prices: Evidence from macroeconometric models. *Journal of Policy Modeling 30*(1), 123–143.

Höppner, F. (2001). A VAR Analysis of the Effects of Fiscal Policy in Germany. Technical report.

Hunt, B. and D. Laxton (2004). Some simulation properties of the major euro area economies in MULTIMOD. *Economic Modelling 21*, 759–783.

Ilzetzki, E., E. G. Mendoza, and C. A. Végh (2011). How Big (Small?) are Fiscal Multipliers? IMF Working Paper WP/11/52.

Ilzetzki, E. and C. A. Végh (2008). Procyclical Fiscal Policy in Developing Countries: Truth or Fiction? NBER working paper 14191.

IMF (2008). World Economic Outlook, October 2008: Financial Stress, Downturns, and Recoveries. Technical report, Washington DC.

IMF (2010). World Economic Outlook October 2010: Recovery, Risk, and Rebalancing. Technical report.

Kirchner, M., J. Cimadomo, and S. Hauptmeier (2010). Transmission of government spending shocks in the Euro Area: time variation and driving forces. European Central Bank working paper series 1219.

Kraay, A. (2012). How large is the government spending multiplier? Evidence from World Bank lending. *Quarterly Journal of Economics 127*(2), 829–887.

Kumhof, M. and D. Laxton (2007). A Party without a Hangover? On the Effects of U.S. Government Deficits. IMF Working Paper WP/07/202.

Kuttner, K. N. and A. S. Posen (2002). Passive Savers and Fiscal Policy Effectiveness in Japan. Technical report.

Leeper, E. M., T. B. Walker, and S.-C. S. Yang (2010). Government investment and fiscal stimulus. *Journal of Monetary Economics 57*(8), 1000–1012.

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.

Mineshima, A., M. Poplawski-Ribeiro, and A. Weber (2013). Size of Fiscal Multipliers. 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.

OECD (2009). OECD Interim Economic Outlook (March 2009). Technical report.

Ortega, E., P. Burriel, J. L. Fernández, E. Ferraz, and S. Hurtado (2007). Update of the Quarterly Model of the Bank of Spain. Documentos de Trabajo 0717.

Pappa, E. (2009). The effects of fiscal expansions: an international comparison. Barcelona Economics Working Paper Series 409.

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

Perotti, R. (2004). Public investment: another (different) look. IGIER working papers 277.

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

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

Ramey, V. A. (2011b). Identifying Government Spending Shocks: It's All in the Timing. *Quarterly Journal of Economics 126*(1), 1–50.

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.

Ratto, M., W. Roeger, and J. i. t. Veld (2009). QUEST III: An estimated open-economy DSGE model of the euro area with fiscal and monetary policy. *Economic Modelling 26*(1), 222–233.

Ravn, M. O., S. Schmitt-Grohé, and M. Uribe (2007). Explaining the Effects of Government Spending Shocks on Consumption and the Real Exchange Rate. NBER working paper 13328.

Rendahl, P. (2012). Fiscal Policy in an Unemployment Crisis. Mimeo.

Roeger, W. and J. i. t. Veld (2004). Some selected simulation experiments with the European commission's QUEST model. *Economic Modelling 21*(5), 785–832.

Roeger, W. and J. i. t. Veld (2010). Fiscal stimulus and exit strategies in the EU: a model-based analysis. European Economy Economic Papers 426.

Romer, C. D. and J. Bernstein (2009). The Job Impact of the American Recovery and Reinvestment Plan. Technical report.

Romer, C. D. and D. H. Romer (1994). What ends recessions? NBER working paper 4765.

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.

Rotemberg, J. J. and M. Woodford (1992). Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity. *Journal of Political Economy 100*(6), 1153–1207.

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.

Shoag, D. (2011). The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns. Mimeo.

Smets, F. and R. Wouters (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association 1*(5), 1123–1175.

Smets, F. and R. Wouters (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review 97*(3), 586–606.

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

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 H. Doucouliagos (2012). *Meta Regression Analysis in Economics and Business*. New York: Routledge.

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.

Taylor, J. B. (1993). *Macroeconomic Policy in a World Economy: From Econometric Design to Practical Operation.* New York: WW Norton.

Tenhofen, J., G. B. Wolff, and K. H. Heppke-Falk (2010). The macroeconomic effects of exogenous fiscal policy shocks in Germany: a disaggregated SVAR analysis. *Jahrbücher für Nationalökonomie und Statistik 230*(3), 328–355.

Truger, A., K. Rietzler, H. Will, and R. Zwiener (2010). Alternative Strategien der Budgetkonsolidierung in österreich nach der Rezession: Gutachten des Instituts für Makroökonomie und Konjunkturforschung (IMK) in der Hans-Böckler-Stiftung im Auftrag der Arbeiterkammer Wien. Technical report, Wien.

Uhlig, H. (2010). Some Fiscal Calculus. *American Economic Review 100*(2), 30–34.

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

Weber, C. E. (1999). Fiscal policy in general equilibrium: empirical estimates from an error correction model. *Applied Economics 31*(7), 907–913.

Wieland, V., T. Cwik, G. J. Müller, S. Schmidt, and M. Wolters (2011). A New Comparative Approach to Macroeconomic Modeling and Policy Analysis. Goethe University of Frankfurt working paper.

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

Zandi, M. (2009). The Economic Impact of the American Recovery and Reinvestment Act. Moody's Economy.com.

Appendix

Table 7 lists included papers, the respective model classes used in these papers and the number of multipliers drawn from these studies.

Table 7: List of included studies

Study	Model Class(es)	# of reported multipliers
Acconcia et al. (2011)	SEE	3
Afonso et al. (2010)	SEE	2
Aiyagari and Christiano (1992)	RBC	1
Alesina and Ardagna (2010)	SEE	14
Alesina and Ardagna (ming)	SEE	24
Almunia et al. (2010)	VAR, SEE	3
Arcangelis and Lamartina (2003)	VAR	8

Continued on next page

Table 7 – cont'd

Study	Model Class(es)	# of rep. mult.
Ardagna (2001)	RBC	5
Auerbach and Gorodnichenko (2012)	VAR	30
Barrell et al. (2004)	DSGE-NK	4
Barrell et al. (2012)	DSGE-NK	72
Barro and Redlick (2011)	SEE	8
Baum and Koester (2011)	VAR	6
Baxter and King (1993)	RBC	5
Bayoumi and Sgherri (2006)	RBC, SEE	13
Beetsma et al. (2006)	VAR	4
Beetsma and Giuliodori (2011)	VAR	10
Bénassy-Quéré and Cimadomo (2006)	VAR	12
Bénétrix and Lane (2009)	VAR	48
Biau and Girard (2005)	VAR	10
Bilbiie et al. (2008)	DSGE-NK, VAR	8
Blanchard and Perotti (2002)	VAR	4
Brückner and Tuladhar (2010)	SEE, VAR	8
Burriel et al. (2010)	VAR	8
Caldara and Kamps (2006)	VAR	10
Caldara and Kamps (2008)	VAR	20
Caldara and Kamps (2012)	VAR	16
Candelon and Lieb (2011)	VAR	12
Canova and Pappa (2007)	VAR	4
Castro Fernández and Hernández Cos (2008)	VAR	17
Congressional Budget Office (2010)	MACRO	8
Christiano et al. (2009)	DSGE-NK	16
Christoffel et al. (2008)	DSGE-NK	1
Church et al. (2000)	MACRO, DSGE-NK	5
Cloyne (2011)	VAR, SEE	11
Coenen and Straub (2005)	DSGE-NK	2
Coenen et al. (2010)	DSGE-NK, MACRO	76
Cogan et al. (2010)	DSGE-NK	4
Corsetti et al. (2009)	DSGE-NK	10
Corsetti et al. (2010)	DSGE-NK	1
Corsetti et al. (2012)	VAR	2
Cwik and Wieland (2011)	DSGE-NK, MACRO	14
Dalsgaard et al. (2001)	MACRO	11
Davig and Leeper (2011)	DSGE-NK	3
Devereux et al. (1996)	RBC	6
Edelberg et al. (1999)	VAR, RBC	3
Eichenbaum and Fisher (2005)	VAR, RBC	3
Elmendorf and Reifschneider (2002)	DSGE-NK	7
Erceg et al. (2006)	DSGE-NK	10
Erceg and Lindé (2010)	DSGE-NK	10
European Commission (2003)	DSGE-NK	8
European Commission (2012)	VAR	32
Fagan et al. (2005)	MACRO	1
Fair (2010)	MACRO	4
Fatás and Mihov (2001)	RBC, VAR	9
Fernandez-Villaverde (2010)	DSGE-NK	2
Fisher and Peters (2010)	VAR	1
Forni et al. (2009)	DSGE-NK	8

Continued on next page

Table 7 – cont'd

Study	Model Class(es)	# of rep. mult.
Freedman et al. (2010)	DSGE-NK	17
Galí et al. (2007)	VAR, DSGE-NK	19
Giordano et al. (2007)	VAR	23
Guajardo et al. (2011)	SEE, VAR	10
Hall (2009)	DSGE-NK, RBC, SEE	8
Heilemann and Findeis (2009)	MACRO	2
Henry et al. (2008)	MACRO, DSGE-NK	33
Höppner (2001)	VAR	6
Hunt and Laxton (2004)	MACRO	5
Ilzetzki and Végh (2008)	VAR	12
Ilzetzki et al. (2011)	VAR	34
IMF (2008)	DSGE-NK, SEE	16
IMF (2010)	DSGE-NK, SEE	14
Kirchner et al. (2010)	VAR	2
Kraay (2012)	SEE	3
Kumhof and Laxton (2007)	DSGE-NK	2
Kuttner and Posen (2002)	VAR	2
Leeper et al. (2010)	RBC	8
Mazraani (2010)	RBC	2
Mountford and Uhlig (2009)	VAR	4
Ortega et al. (2007)	MACRO	1
Pappa (2009)	VAR	15
Perotti (2004)	VAR	10
Perotti (2005)	VAR	30
Ramey and Shapiro (1998)	RBC	2
Ramey (2011b)	VAR	6
Ratto et al. (2009)	DSGE-NK	2
Ravn et al. (2007)	RBC, VAR	4
Rendahl (2012)	DSGE-NK	2
Roeger and Veld (2004)	DSGE-NK	23
Romer and Romer (1994)	SEE, MACRO	2
Romer and Bernstein (2009)	DSGE-NK	2
Romer and Romer (2010)	SEE, VAR	6
Rotemberg and Woodford (1992)	RBC, VAR	3
Shoag (2011)	SEE	7
Smets and Wouters (2003)	DSGE-NK	1
Smets and Wouters (2007)	DSGE-NK	1
Taylor (1993)	DSGE-NK	14
Tenhofen et al. (2010)	VAR	5
Truger et al. (2010)	MACRO	5
Uhlig (2010)	RBC	4
van Brusselen (2009)	MACRO	26
Weber (1999)	SEE	11
Wieland et al. (2011)	DSGE-NK	3
Woodford (2011)	DSGE-NK	1
Zandi (2009)	MACRO	4

Table 8: Robustness – stepwise exclusion of comprehensive studies (Dep. Var.: multiplier)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
κ	0.7274***	0.7188***	0.7418***	0.7348***	0.7331***	0.7193***	0.6904***
	(3.089)	(2.929)	(5.019)	(3.013)	(3.645)	(3.426)	(3.400)
fiscal impulse							
CONS	0.1122	0.04424	0.1108	0.1281	0.1024	0.1317	0.1148
	(0.9746)	(0.3904)	(0.8869)	(1.094)	(0.7052)	(1.053)	(1.002)
INVEST	0.6169***	0.6608***	0.6235***	0.6294***	0.6130***	0.4880***	0.6161***
	(3.566)	(3.357)	(3.304)	(3.658)	(3.544)	(3.250)	(3.559)
MILIT	-0.2297	-0.2345	-0.2407	-0.2188	-0.2315	-0.2520	-0.2325
	(-1.418)	(-1.436)	(-1.506)	(-1.342)	(-1.429)	(-1.573)	(-1.428)
TAX	-0.3208***	-0.3234***	-0.3190***	-0.2976***	-0.3175***	-0.3352***	-0.3225***
	(-3.200)	(-3.047)	(-3.153)	(-2.760)	(-2.910)	(-3.303)	(-3.219)
TRANS	-0.3852***	-0.3464**	-0.3820***	-0.3709***	-0.3878***	-0.4217***	-0.3863***
	(-3.552)	(-2.435)	(-3.509)	(-3.336)	(-3.585)	(-3.949)	(-3.564)
EMPLOY	0.2348	0.2260	0.2339	0.2491	0.2316	0.2141	0.2334
	(1.429)	(1.376)	(1.410)	(1.506)	(1.383)	(1.292)	(1.418)
DEFICIT	-0.1374	-0.1418	-0.1346	-0.1328	-0.1382	-0.1489	-0.1364
	(-1.171)	(-1.204)	(-1.150)	(-1.090)	(-1.177)	(-1.299)	(-1.160)
model class							
RBC	-0.5848**	-0.5814**	-0.5981**	-0.5835**	-0.5840**	-0.5871**	-0.5830**
	(-2.277)	(-2.247)	(-2.327)	(-2.270)	(-2.265)	(-2.285)	(-2.270)
DSGE-NK	0.1888	0.1805	0.1733	0.1902	0.1864	0.1997	0.1884
	(0.9513)	(0.9005)	(0.8750)	(0.9516)	(0.9325)	(1.004)	(0.9423)
MACRO	0.4287*	0.4176*	0.4129*	0.2578	0.4239*	0.4397*	0.4440*
	(1.762)	(1.702)	(1.698)	(0.6338)	(1.728)	(1.809)	(1.871)
SEE	0.08541	0.09286	0.07814	0.07843	0.08807	0.1004	0.1118
	(0.4308)	(0.4648)	(0.3899)	(0.3950)	(0.4412)	(0.4945)	(0.5716)
control variables							
PEAK	0.3362**	0.3342**	0.3382**	0.3296**	0.3359**	0.3624**	0.3711***
	(2.477)	(2.458)	(2.389)	(2.453)	(2.463)	(2.446)	(2.768)
HORIZON	0.1670***	0.1683***	0.1665***	0.1655***	0.1671***	0.1627***	0.1485***
	(4.623)	(4.613)	(4.515)	(4.501)	(4.585)	(3.474)	(4.337)
HORIZON^2	-0.002467	-0.003473	-0.004788	-0.004947	-0.002600	-0.0008516	0.02527
	(-0.06117)	(-0.08564)	(-0.1198)	(-0.1251)	(-0.06401)	(-0.01852)	(0.7299)
PEAK*HOR	0.01155	0.01421	-0.04445	0.01054	0.01123	0.009948	0.03443
	(0.1229)	(0.1489)	(-0.4583)	(0.1120)	(0.1190)	(0.1001)	(0.3735)
PEAK*HOR^2	0.08584	0.09070	0.05818	0.08843	0.08599	0.07074	0.05650
	(1.167)	(1.230)	(0.8018)	(1.213)	(1.157)	(0.9394)	(0.8141)
M/GDP	-0.01033***	-0.01197***	-0.01004***	-0.01122***	-0.01129***	-0.01016***	-0.01126***
	(-3.301)	(-3.359)	(-2.931)	(-3.117)	(-2.276)	(-3.267)	(-3.529)
N	1063	987	1029	1030	991	1015	1031
Adj.R^2	0.4138	0.4047	0.4163	0.4099	0.3904	0.3861	0.4223
ℓ	-882.7	-845.1	-854.9	-867.5	-855.0	-844.7	-845.5

reference: SPEND, VAR, CUM

Column-wise excluded studies: (1) total sample; (2) Coenen et al. (2010); (3) Ilzetzki et al. (2011); (4) Henry et al. (2008); (5) Barrell et al. (2012); (6) Bénétrix and Lane (2009); (7) European Commission (2012)

*, **, *** indicate significance at the 10, 5, 1 percent level respectively, Standard errors in parentheses

Table 9: Robustness – weighted samples (Dep. Var.: weighted multiplier)

	(1) total[a]		(2) subsample I[b]		(3) subsample II[c]
κ	0.06686***		0.06880**		0.09915***
	(2.692)		(2.188)		(6.834)
fiscal impulse					
CONS	0.6364		0.2858		0.8868**
	(1.592)		(0.9747)		(2.198)
INVEST	0.3483		0.8505*		-0.05612
	(1.235)		(1.843)		(-0.2359)
MILIT	0.1432		-1.328**		1.729***
	(0.5449)		(-2.094)		(3.090)
TAX	-0.1803		-0.04280		-0.5537***
	(-1.154)		(-0.2094)		(-5.450)
TRANS	-0.3515		-0.1954		-0.5656**
	(-1.531)		(-0.7786)		(-2.583)
EMPLOY	-0.09043		0.8962***		-0.2601
	(-0.4542)		(3.214)		(-1.240)
DEFICIT	-0.05273		0.09228		
	(-0.3085)		(0.5641)		
model class					
RBC	-0.2995*				-0.2394
	(-1.713)				(-0.9305)
DSGE-NK	0.1603				
	(0.5734)				
MACRO	0.3386				-0.3542
	(1.295)				(-0.9256)
SEE	0.05390		1.043**		
	(0.2538)		(2.357)		
control variables					
PEAK	0.1468		0.6937**		0.01483
	(0.7686)		(2.306)		(0.08176)
HORIZON	0.07630*		0.1075		0.007719
	(1.726)		(1.392)		(0.1352)
$HORIZON^2$	-0.01963		0.03639		-0.0003643
	(-0.5425)		(0.3488)		(-0.007007)
PEAK*HOR	0.2219		0.3070		0.2463
	(0.8291)		(0.9102)		(1.505)
$PEAK*HOR^2$	0.2441		-0.009710		0.2068
	(1.461)		(-0.05463)		(1.207)
M/GDP	-0.01882***		-0.02087*		-0.02557***
	(-3.387)		(-1.777)		(-2.976)
additional characteristics					
		LOGOBS	0.1765	NONRICARD	0.01008**
			(0.8356)		(2.224)
		BP	0.0006337	LOANABLE	-1.105**
			(0.1295)		(-2.237)
		RA	-0.1438*	FIXREAL	-0.2260
			(-1.850)		(-1.476)
		SR	-0.4270***	ZLB	0.8986***
			(-9.068)		(5.138)
		WAR	-0.4922**	CLOSED	0.5600***
			(-2.379)		(5.273)
		IV	-0.3419		
			(-1.013)		
		CA	-0.8932***		
			(-7.344)		
		SEE*WAR	1.767***		
			(3.174)		
		ANNUAL	-0.5674**		
			(-2.233)		
		AVYEAR	-0.02435***		
			(-4.644)		
N	1063		556		504
$Adj.R^2$	0.8718		0.8846		0.9529
ℓ	1484		836.8		861.3

[a] reference: SPEND, VAR, CUM
[b] reference: SPEND, VAR, CUM, NAR, QUART
[c] reference: SPEND, DSGE-NK, CUM, INFTARG, OPEN
*, **, *** indicate significance at the 10, 5, 1 percent level respectively, Standard errors in parentheses

Table 10: Robustness – median values; probit and logit model (Dep. Var.: multiplier)

	(1) median	(2) probit[a]	(3) logit[a]
κ	1.029***	0.7069***	1.301***
	(8.635)	(2.601)	(2.701)
fiscal impulse			
CONS	-0.1529	0.2740	0.4373
	(-0.8934)	(0.9663)	(0.9025)
INVEST	-0.2349	0.9783***	1.651***
	(-1.352)	(3.835)	(3.763)
MILIT	0.02362	-0.8721	-1.521
	(0.1005)	(-1.368)	(-1.317)
TAX	-0.3179**	-1.059***	-1.980***
	(-2.094)	(-3.845)	(-3.902)
TRANS	-0.7028***	-1.764***	-3.148***
	(-3.861)	(-3.777)	(-3.698)
EMPLOY	-0.3408	0.9639*	1.555
	(-1.572)	(1.719)	(1.567)
DEFICIT	-0.4928	-1.288	-2.122
	(-1.532)	(-1.555)	(-1.054)
model class			
RBC	-0.3367*	-1.338	-2.674
	(-1.759)	(-1.549)	(-1.482)
DSGE-NK	-0.04699	0.4260	0.8355*
	(-0.3069)	(1.456)	(1.731)
MACRO	0.3718***	1.318	3.010*
	(2.798)	(1.507)	(1.814)
SEE	-0.09193	-0.5209	-0.7213
	(-0.3561)	(-0.8643)	(-0.6526)
control variables			
PEAK	-0.09359	0.5035***	0.8799***
	(-0.7958)	(2.002)	(2.014)
HORIZON	0.05631	0.3577***	0.6074***
	(0.5530)	(3.999)	(3.986)
HORIZON^2	0.02744	-0.07880	-0.1630
	(0.4197)	(-0.7443)	(-0.9003)
PEAK*HOR	0.02725	0.3582	0.6513
	(0.1961)	(1.309)	(1.346)
PEAK*HOR^2	0.03016	0.4568***	0.8117***
	(0.3473)	(2.697)	(2.707)
M/GDP	-0.02084***	-0.02672**	-0.04812**
	(-2.939)	(-2.082)	(-2.065)
N	104	943	943
$Adj.R^2$	0.1597	0.3363	0.3407
ℓ	-68.97	-419.8	-417

reference: SPEND, VAR, CUM

*, **, *** 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.