

# 18 BILLION AT ONE BLOW - EVALUATING GERMANY'S TWENTY BIGGEST TAX EXPENDITURES

Michael Thöne, University of Cologne\*

## INTRODUCTION

*“Periodically evaluating the size and effectiveness of tax expenditures is a necessary (although not sufficient) requirement for good government.”*  
(Burman, 2003)

IN JULY 2007 THE GERMAN FEDERAL MINISTRY OF Finance commissioned the *Centre for European Economic Research (ZEW)* in Mannheim, the Danish consultancy firm *Copenhagen Economics* and, as team leader, the *FiFo Institute for Public Economics at the University of Cologne (FiFo Köln)* with an encompassing evaluation of the twenty biggest tax expenditures in Germany.

The project covers 85 percent of the fiscal volume of all German tax subsidies.<sup>1</sup> For the first time, so many big tax subsidies have been evaluated at one blow. Two years later, and with the work of twelve researchers, three volumes of evaluation reports with 620 pages had been produced. With 18 billion euro in tax subsidies covered, this was the biggest evaluation of tax expenditures ever concluded in Germany, and one of the biggest worldwide. The analysis covered exemptions from corporate and personal income taxes, value added tax (VAT), and energy taxes.

In the past 15 years, the relative importance of direct subsidies and tax subsidies in Germany has changed markedly. In 1995, federal direct subsidies amounted to 0.5 percent of GDP, and federal tax subsidies amounted 0.5 percent of GDP, as well. In 2010, direct outlay measures had fallen to 0.28 percent of GDP, whereas tax subsidies had risen to 0.75 percent.

---

\*The author would like to thank Sigurd Naess-Schmidt, Friedrich Heinemann, Anna Jung, Eva Gerhards, Christina Elschner, Robert Schwager, Margit Kraus, Christian Bergs, Thilo Schaefer, and Christoph Spengel for co-working on the evaluation. Thanks to Christian Kastrop, Ralf Debelius, Arno Diekmann, and the BMF-Team, thanks to numerous discussants within the German Federal Administration and the Federal Audit Office, thanks to Clemens Fuest, Rita de la Feria, and the participants of the Cologne-Oxford Tax Workshop at the Oxford University Centre for Business Taxation 2009; and thanks to Lori Metcalf, G. Thomas Woodward, Thomas Hungerford, and Leonard Burman for helpful comments on an earlier version presented at the National Tax Association's 104th Annual Conference on Taxation, Nov 17, 2011, in New Orleans. Finally thanks to Anna Rauch for editorial assistance on this paper. All remaining mistakes are mine.

Today, tax subsidies on the federal level outnumber direct subsidies by a factor of almost three.

For many years, this relative shift in the instrumentation of transfer policies was not accompanied by a corresponding shift of focus in the governance of subsidies. While direct subsidies were evaluated rather regularly, tax expenditures went almost unnoticed. This has changed with our evaluation of the twenty biggest tax expenditures.

This paper gives an overview of the evaluation project and its results. The remainder is organized as follows: first, we outline the uniform evaluation framework and discuss its role in face of very dissimilar tax expenditures (section 2); then we explain the different steps of the evaluations according to the framework (section 3); finally, we give a brief overview of the actual evaluation results and draw some conclusions (section 4).

## THE UNIFORM EVALUATION FRAMEWORK

### How Much Standardization is Possible in the Evaluation of the Dissimilar German Tax Expenditures?

The joint scheme employed in the evaluations of the different tax expenditures is based on a general scheme for the control of subsidies developed by Thöne (2003) for the German Federal Ministry of Finance. The basic task of the common evaluation scheme for different and dissimilar tax expenditures is to warrant comparability as far as comparability makes sense. The common framework has several functions:

- It aims to ensure that evaluations are based on a transparent understanding of the role of subsidies and tax expenditures in the market economy.
- It ensures that all tax expenditures are subject to the same questions, thus instituting equivalent quality standards for the evaluations.
- It should, as far as possible, aim to establish comparability in the economic indicators and benchmarks used.

But the framework must also disclose the fundamental limits of such standardization. A common framework with the same criteria used for all tax expenditures can only provide added value if the individuality of the policy objectives and the conditions of their implementation are respected. This tension between individuality and comparability is especially high in the evaluation of the biggest German tax expenditures. In many respects the individuality must dominate, especially in view of the different taxes concerned and the diverse subsidy objectives.

The similarities of the tax expenditures examined are limited primarily because they are enacted in quite different taxes. This is due to the size of the German coverage of tax subsidies (Appendix 2 of the official Subsidy Report) and “other tax rules” (Appendix 3 of the Subsidy Report). The biannual subsidy report of the Federal Government of Germany gives an overview of tax expenditures which are subject to the federal legislature. In practice, in the intertwined German tax system, *all* taxes with sizeable revenue fall under federal legislature. Tax expenditures from the taxes on private and corporate income, from general consumption tax (VAT) as well as from excise taxes and the like are covered. Since the 20 biggest tax expenditures actually reflected a good part of this spectrum, many different tax benchmarks had to be referred to. Thus, the common framework cannot prove uniform criteria to distinguish between the tax rule and the exemption from this rule, i.e. the tax expenditure.

Also, the tax expenditures examined differ strongly in terms of policy objectives and areas of application. Some serve as “traditional” sectoral or regional subsidies with a primary job goal. Some are awarded in the field of environmental and energy policy. Others pursue ‘merit wants’ and (liberal) paternalistic purposes, encouraging people to consume more culture and arts or to save more for retirement or housing. Then there are tax credits to compensate for perceived disadvantages of certain groups of taxpayers. Finally, one of the tax reliefs examined was even identified as a specific benefit, which unlike any other is not meant to produce any behavioral response.<sup>2</sup> As a consequence, the variety of purposes and goals prevents common standards and measurement methods from being defined in order to identify the success or failure of different tax expenditures.

In practice, these two problems relate especially to the third bullet point above. Establishing comparability of the economic indicators and benchmarks of success is feasible only to a very limited extent. Accordingly, the common framework of tax expenditure evaluation mainly relates to the first two points above.

### The Structure of Each Evaluation

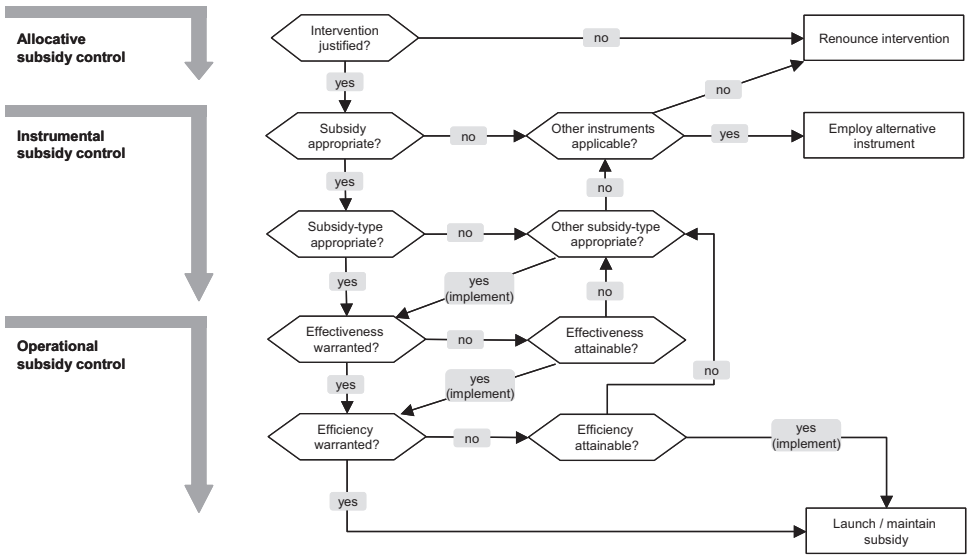
Each subsidy is unique and calls for a unique (i.e. tailor-made) evaluation. This is also true for tax expenditures. Tax credits to promote research and development in the corporate sector have very little in common with reduced consumption tax rates for merit goods like newspapers, food, and baby clothing. These two exemplary tax expenditures differ with regard to their basic justification, to their instrumental design, to their potential output and outcome dimensions, and to the methods of calculating size and effects of the measures.

Under these circumstances, coherence and consistency among evaluations relies heavily on the comparability of methods. Here, the most important feature is an equal structure of the individual evaluations. Each evaluation must ask the same questions. These are not necessarily identical questions, but questions with an identical function within the common framework. Each evaluation conforms to the following blueprint:

1. Brief description of the tax expenditure and its evolution over time.
2. Measurement of the actual volume of the tax expenditure.
3. Record of past evaluations and findings from academic research.
4. Core Evaluation:
  - a. Transparency of the measure
  - b. Rationale of the subsidy
  - c. Relevance of the subsidy and instrumental subsidy-control
  - d. Testing for effectiveness
  - e. Testing for efficiency
5. Conclusions and proposals for actions to be taken

The core evaluation (step 4) is modeled after the scheme for an optimal subsidy control, which was originally developed by FiFo Köln and subsequently adopted by the Federal Ministry of Finance (figure 2).<sup>3</sup>

Figure 1: Scheme of Optimal Subsidy Control



© Thöne (2003).

In general, this scheme is rather straightforward; it does not strive for originality. Rather, the scheme is the result of an international benchmarking endeavor; it reflects the notion of a common best practice. The framework poses the basic questions every economist would ask: Is the subsidy justified? Is it relevant? Is it well-designed? Is it effective? Is it efficient?

In the course of this paper, these questions will be further elaborated for tax expenditures.

**THE STEPS OF THE EVALUATIONS**

In this chapter, we briefly characterize the individual steps for the evaluations, as implemented according to the uniform framework.

**Short Subsidy Description and History**

Each evaluation report begins with a description of the tax rules to be examined. This description provides information on aspects such as: legal and technical design of the tax expenditure, purpose of the tax expenditure according to the latest official statement, initial rationale (if different), date of the first introduction of the tax expenditure and timing of significant changes in the meantime, and

instrumental environment (e.g. parallel measures with similar purposes).

Not all points are relevant for all tax expenditures evaluated. For example, the older a tax break is (the oldest measure evaluated is from 1930), the more important the analysis of past justification patterns and the genesis of the instrument.

**Measurement of the Actual Volume of Tax Expenditures**

The measurement of the volume of tax expenditures is a central question in the literature on this subsidy form. As opposed to direct subsidies, the fiscal costs of tax expenditures are not transparent from the outset. Depending on the method used and on available databases, the quality and accuracy of the estimates vary strongly.

The estimate of the fiscal volume is important in two ways: For each individual measure, and, independently, for the sum of all tax expenditures. First, the quantification is central to the individual evaluation of each tax expenditure. The loss of tax revenue is the price that is paid for the behavioral response of the beneficiaries. Where an analysis of cost-efficiency is possible, this figure is the denominator in the calculation. Conversely, any

analysis of tax expenditures seems pointless without quantifying the fiscal relevance.

Second, the total volume of all tax expenditures receives special attention since this figure is a policy parameter itself. The total volume of grants and tax breaks is an obvious indicator to judge the success of government in the reduction of subsidies. But it is very important to be able to clearly identify the causes of changes in financial volume. Especially, improvements in *measuring* tax expenditures should not be confused with factual changes in total volume. An example from Germany's 21<sup>st</sup> Subsidy Report (2007) was the first time quantification of the "tonnage tax." The half billion euro, to which this tax expenditure was estimated by the administration, has no effect on the actual change in the volume of subsidies because the measure has existed since 1999. It only becomes visible for the first time in a specific quantity.

The bulk of the tax benefits which are listed in the Subsidy Report are quantified (82 percent at the time of the 21<sup>st</sup> Subsidy Report). Of the so called "other tax rules" (Annex 3 of the Subsidy Report), which after 1977 are no longer called a tax subsidy, 21 of 53 measures (i.e. 40 percent) are quantified.

In general, the research on tax expenditures traditionally suffers from a lack of internationally comparable standard definitions. To some extent this is the case because national tax systems are individual in many aspects. And if the tax rules are not uniform, the exceptions to the rule – i.e. the tax expenditures cannot be defined uniformly as well.

Yet a standard definition is not necessary in the context of a national evaluation. For us, two points with regard to the quantification of the German tax expenditure are particularly important:

1. Tax expenditures are the proverbial exception to the rule. Identifying an unambiguous benchmark tax to which the measure in question as the exception can be tricky.
2. In general, there are three different methods to calculate the volume of tax expenditures. In the evaluation, we used the standard method (revenue forgone) but supplemented it with two more approaches to gain deeper insight in the full costs of the measures.

### **Benchmarking the Tax System**

The question of the best benchmark tax system cannot be answered objectively. The clearer and

the more timeless the tax rule is, the easier the exception to this rule can be identified. In practice, however, taxes do not comply with this platonic ideal. They reflect the democratic decision-making process: they are complex and often contradictory, they change over time, and they are controversial up to their normative foundations. In such a living tax system the tax benchmarks are not self-evident, they must be developed meticulously.

We illustrate this challenging task with the income tax treatment of private pension plans and their funding. In the field of personal income taxation (PIT), Germany traditionally adheres to comprehensive income tax of the Schanz-Haig-Simons type (SHS). In practice, this has never been implemented in a very stringent manner: Germany has given up taxing the imputed income from owner-occupied housing in 1987. Also, Germany displays an almost traditional weakness in capital gains taxation. Accordingly, the German tax model can be characterized as a "low-key" comprehensive income tax. For the last 10 years the German income tax developed even further away from the basic model so that the determination of the benchmark income tax has become even more difficult.

In 2002, the tax subsidization of privately funded pensions was introduced. Among other measures, individual contributions to pension funds can be made out of untaxed income; the income tax is levied decades later when the pensions are paid out. An analogous change is implemented gradually for obligatory retirement benefits and pensions. Thus, for certain income types, Germany introduced an expenditure tax instead of the comprehensive income tax.<sup>4</sup> In addition, operative from 2009, Germany also has implemented a tax reform that encompasses core elements of a Dual income tax (DIT), though – officially – it is not referred to as such. Though it is mostly an issue of corporate income taxation (CIT), it also extends directly to personal income tax (PIT): interest income, dividends, and capital gains on shares, or other financial instruments realized by individuals as nonbusiness income, are subject to a flat tax rate of 25 percent instead of the usual progressive rate that went up to 45 percent. This tax is levied as a withholding tax. Taking both reforms together, we can characterize the German personal income tax not as dual, but as triple income tax (figure 2).

Figure 2: Triple Tax Benchmark in German Income Tax

Income from...	thereof...	Tax regime		
		Standard Comprehensive PIT (SHS-Income)	Flat Rate Withholding Tax (Dual Income Tax)	Expenditure Tax (Consumed Income)
1. Agriculture and Forestry		X		
2. Trade and Business		X		
3. Self-employment		X		
4. Dependent labor		X		
5. Capital income	up to 800 €			X
	from 801 €		X	
6. Rent and Leasing		X		
7. Other income	Old age pensions, public pensions *			X
	Subsidized Riestler-/Rürup-Pensions			X
	Capital gains		X	
	Remaining other income	X		

\* Gradual introduction until 2040.

Effectively, this momentary portrait of the “triple” German income tax was used as the benchmark in the evaluation of tax expenditures in the income tax. Yet, from a systematic perspective, this decision remains ambivalent. Because the benchmark reflects the current principles of German PIT, it is the politically most relevant benchmark. Yet, it is also a benchmark which allows few intertemporal comparisons let alone, international comparisons.

Similar experiences have been made elsewhere. As a consequence, for example, the budget proposal of the President of the United States for many years has calculated tax expenditures in several different definitions and presented them to Congress. The climax of this development was reached in Fiscal Year 2007 when four different benchmarks were used to define tax expenditures.<sup>5</sup> This approach is more systematic and more honest. On the other hand, the diversity may also add to the public confusion around tax expenditures.

The implicit conflict between clarity and political relevance cannot be solved without trade-offs. With tax systems increasingly reflecting the complexities of life in a globalized economy, problems of benchmarking taxes will probably even increase.

### Measuring the Costs of Tax Expenditures

The literature discusses three main methods to calculate the fiscal volume of tax expenditures:

revenue foregone, revenue gain, and outlay equivalent. The revenue gain method and the outlay equivalent method each provide more information, but are more complicated and costly to implement. This is justified in individual cases. But when it comes to calculating the fiscal volumes of many tax expenditures, revenue foregone is the method of choice. Consequently, it is recommended by the Organisation for Economic Co-operation and Development (OECD) as best practice (OECD, 2004).

The German administration uses the revenue forgone method as well. In our evaluations, the official quantifications were typically subjected to critical analysis; in some cases, the official quantification was compared with a separate calculation. In addition, selected evaluations deepen the analysis of the economic cost of tax expenditure in two ways. First, we made exemplary use of the concept of marginal costs of public funds (Dahlby, 2008). It provides an easy way to calculate the overall economic shadow cost of public expenditures and foregone revenue, respectively. For a tax subsidy, the deadweight loss arises from the fact that the rates of the respective tax (or another, more politically more probable tax) must be higher than without the subsidy. The use of such shadow prices of public spending is very straightforward for the evaluation of public policies. The MCF-factor we calculated for our treatment was in the normal range (1.2 to 1.3).

Second, the granting of tax expenditures can cause considerable administrative effort that should be added to the total cost of operation. The standard approaches to calculate the volume of tax expenditures do not extend to these costs. Where possible, we carried out estimates for administrative costs by extending the Federal Government's Standard Cost Model<sup>6</sup> to tax expenditures. If annual case numbers and information on the approximate administrative effort per case are available, standard cost models provide an easy and straightforward way to assess these potentially decisive costs.

### **Transparency of the Measure**

The essential role of budget transparency for political accountability and for democratic control is undisputed now. Many elements of modern budget reforms are primarily aimed at increasing transparency. This reasoning should also extend to the transparency of tax expenditures.

In the evaluation of a tax subsidy, transparency is a prerequisite, as well as a result, of the undertaking; the more transparent the measure is, especially in user-statistics and accompanying administrative knowledge, the more thorough and informative the evaluation can become. In the next step, the evaluation will further enhance the transparency of the measure by deepening the understanding of instrumental performance, subsidy effects, etc. Thus, subsidy transparency reinforces itself.

In the actual evaluation of German tax expenditures, the transparency concerns started at a rather elementary level. The best example to illustrate this fact is the almost total lack of information on the evaluated employees' savings tax credit. Apart from data on total fiscal volume, no systematic information whatsoever (e.g. on the number of recipients or on the structure of the savings funded) has ever been given to the public. Even the evaluators had no access to these data, since they are "owned" by the states (who are still responsible for income tax administration), not by federal government. Obviously, this is not acceptable for a measure that has been in force for 50 years now and has incurred fiscal costs of 5 billion euro in only the last 20 years.

### **Rationale of the Tax Measure**

As with any appraisal of a government measure, the focus of the evaluation of tax expenditure lies on the benefit of return that is achieved for taxpayers' money. Here, the crucial question is whose

value is relevant. Depending on the answer, the success of the measure may be viewed quite differently. For obvious reasons, the financial value of a subsidy to the recipient cannot be used here. The same is true for the political value to the institutions granting the tax expenditure. The ultimate yardstick for any appraisal of a tax subsidy is the public interest.

The public interest may, but does not necessarily have to, coincide with the statutory purpose of a tax expenditure. Many subsidies may indeed be the most effective instrument to serve the common good. Yet, the dubious reputation of subsidies comes from the fact they are also suited very well to serve vested interests with little relation to, or even against, the public interest. Thus, the objectives of a specific tax expenditure, as defined by law, cannot always be used as the ultimate standard for an evaluation. Still, the objectives intended by the legislature are the starting point for any evaluation. They serve as the first (and in many cases also the most important), but not as a final benchmark for the evaluation of the social and economic benefit of a tax expenditure.

In the evaluation of German tax expenditures, we used a three-tier benchmark. First, we examined whether the observed tax expenditures reached their own statutory purposes. The second tier is to assess whether they comply with the general guidelines for subsidy policy of the Federal Government as proclaimed in March 2006. This commitment of the Federal Government provides additional quality requirements for subsidies. Here, tax expenditures are viewed rather critically.

Thirdly, any tax subsidy is examined in terms of the public interest. Obviously, this is not an easy task. The "common good" and "public welfare" are vague concepts. All attempts to operationalize them are necessarily tainted by value-judgment to some degree. In the evaluation, this problem could not be solved, but it was effectively contained by using simple and plausible principles of economic efficiency and of horizontal and vertical equity.

These three levels of analysis do not necessarily have to provide different results. The differences are mainly relevant when meeting the grant objectives of the legislature alone cannot suffice as a yardstick for evaluation.

### **Instrumental Subsidy Control**

The instrumental subsidy control mainly asks one question: is the measure well designed to



achieve the desired objectives? If not, can improvements in the instrumentation provide better results?

To some degree, the suitability of the instrument is an issue that might be best judged after empirically testing its effectiveness, i.e. after the next step of the common evaluation framework. This implies, however, that in practice there is room to experiment what type of subsidy could achieve the best results for a particular purpose. This may well be the Popper's ideal of piecemeal social engineering. Yet, in practice, usually this room is not given. Accordingly, the evaluation scheme poses the instrumental question from the perspective of the planner – thus evaluating the decisions made in the creation of the tax expenditure. This sequence, of course, does not exclude feeding empirical findings from the final step back into the instrumental subsidy control.

The instrumental subsidy control asks two questions in a row: Should the planned government intervention take the form of a tax expenditure? If so, what form should that tax expenditure take precisely?

When government intervention is justified in a particular problem in principle, in addition to tax incentives, a number of other instruments are available: direct subsidies/transfers, low-interest loans, government loan guarantees, information campaigns, awards, or (on the other side) command and control instruments. Compared to these, tax expenditures must be proven in the first step as the instrument that is most suitable to reach the respective target of intervention. Here, mixed instrumental strategies may also be considered.

If this step affirms that the intervention should take the form of a tax subsidy, the best design of the tax expenditure has to be established. The first task is choice of the best tax. Often, only one tax comes into question. But in other contexts, this question is quite open. For example, fiscal instruments to fight moonlighting and to advance legal low-skill jobs are attached to the value added tax in many European Countries; in Germany the objective is pursued income tax credits. Thus, in the instrumental control of the German income tax credit for craftsmen services and for labor-intensive services, the choice of the income tax of the VAT was reviewed (and, in the end, affirmed).

Once the appropriate tax is identified, the question arises whether the tax expenditure in question has taken the most appropriate form as tax exemption, tax allowance, tax credit, tax rate relief, or tax defer-

ral. Beyond these basic distinctions, the instrument design was reviewed in great detail. The analysis conducted for the employee savings tax credit gives a good example. This tax expenditure has been used for 50 years now; and it has been changed again and again. Also, the statutory purposes have changed with the times. Today, they are no longer clearly identifiable. In the course of the evaluation, eight different purposes have been identified that have been pursued temporarily or continuously with this tax expenditure. We confronted these potential purposes with the two most important design-elements of this tax credit, the group of eligible beneficiaries and the group of eligible investments. Then, for each of the possible policy objectives, optimal instrument structure in the two central dimensions was formulated and compared with the current design. In this specific example, we come to the conclusion that the employee savings tax credit is ill-designed for all potential targets either it is too wide or, more often, too narrow.

#### **Operational Subsidy Control: Effectiveness and Efficiency of the Measure**

The operative subsidy control — the review of the effectiveness and efficiency of each tax expenditure — is the focal point of each evaluation. Yet, there are only a few common points to be reported. This paradoxical statement is explained by a two-fold diversity: the diversity of targets and subsidy purposes of the big German tax expenditure on the one hand, and the variety of methods that can be employed in the impact analysis on the other hand. Both dimensions are difficult to be lumped together. Basically, operative subsidy control includes two elements:

- Effectiveness: To what extent did the tax subsidy reach its objectives? Was the level of target attainment acceptable? Can a level of effectiveness found to be insufficient be improved through viable reforms?
- Efficiency: What is the cost per unit of target attainment? Is the observed cost-effectiveness ratio acceptable? Can the cost per unit of subsidy-output (or subsidy-outcome) be reduced by improving the instrument and its governance?

In practice, the effectiveness test clearly dominates the efficiency review. Central for both aspects

is the question of the relevant target. According to which standards should the achievements with regard to effectiveness and efficiency be measured? Ideally, this should be the overarching objectives for which the tax expenditure is granted. Yet such an outcome orientation can only very rarely lead to reliable and valid empirical results. Usually the successes and failures of a subsidy program can only be determined on the intermediate target level of direct program outputs.

Beyond these few basic notions, a wide variety of methods to analyze the effects of tax expenditure exists. Original empirical research would always be desirable, but the cost of conducting surveys and collecting data easily becomes prohibitive. For tax expenditures, empirical research is even more difficult because of tax secrecy.

In the actual evaluations of the 20 biggest tax expenditures, a method mix has been used that, in general, combines three different elements: model-based simulations, either based on extensive micro data sets or calibrated with empirical parameters; econometric estimates of the effects on the basis of existing databases and sources; meta-analysis of existing empirical analyses.

At the end, to the surprise of the evaluators, the operative efficiency, in practice, has developed into a step of analysis with low importance. The reason for this may lie in the high difficulty of effective analysis and the concurrent efficiency check. Quantifying the effects of tax expenditure reliably has evolved into the main challenge of the evaluation.

Where this is successful, an efficiency ratio can easily be calculated. Yet due to the great diversity of analyzed subsidy purposes, the use of standard indicators (subsidy effect per euro) convey very little additional information since they are not comparable. Because of these constraints, the operative efficiency ranks rather low in the evaluation reports. This would be different if many subsidies with comparable goals would be evaluated. With such thematic or sectoral approaches, efficiency indicators could play a larger role.

## RESULTS

### Overview of Evaluation Results

After summarizing the common methodology of the evaluation, we present the results of the

evaluations in table form with traffic lights. This is the shortest of three different reporting formats representative of the evaluation results. The assessments are summarized according to a uniform, deliberately striking pattern in the form of the three colors of traffic lights:

- Green light: The tax expenditure can be maintained as long as the subsidy target is pursued. This recommendation does not preclude minor improvements or the advice to evaluate the measure at regular intervals.
- Yellow light: The tax expenditure should be heavily revised, as it has done badly at least according to one of the evaluation criteria. Often, the warning yellow light goes hand in hand with the recommendation to deepen the evaluation and to check again as soon as better data are available.
- Red light: The tax expenditure should be abolished. Deficiencies in one or several evaluation criteria are so serious that they cannot be resolved by adjustments and reforms.

In table 1 the 20 tax expenditures evaluated are ordered according to their fiscal volume.

Altogether, tax credits and reductions with an estimated volume of 18.12 billion euro for 2008 were evaluated. That amounted to 85 percent of the total sum of tax expenditures. The overall picture of the evaluated results is mixed. Of the 20 measures investigated, five received a red traffic light (25 percent), ten a yellow (50 percent) and five the green light (25 percent). The dominance of urgent need for revision (yellow traffic light) is even more apparent when looking at the financial volume of the three groups; the group with yellow traffic light comprises more than 10.5 billion euro, or 58 percent of the total volume. In the red group the revenue losses amount to 26.5 percent of the total, affirming its approximate 25 percent share. The group of tax expenditures rated green is the smallest. Together they comprise 2.8 billion euro, but only 15.5 percent of the revenue loss of all twenty big tax expenditures.

## CONCLUSIONS BEYOND THE INDIVIDUAL EVALUATIONS

Apart from the immediate results for the individual tax expenditures, our bundled, and in this



*Table 1*  
**Results of Twenty Evaluations**

	Act	Tax	Purpose of Tax Expenditure	Million euro	Traffic Light
1	§ 35a Abs. 2 Satz 2 EStG	PIT	Income tax credit for craftsmen-servives	2 185	yellow
2	§ 3b EStG	PIT	Tax free bonuses for night and weekend work	2 000	red
3	§ 9 Abs. 3 StromStG	Energy Tax	Ecotax reduction for many industrial companies	1 850	yellow
4	§ 12 Abs. 2 Nr. 1 und 2	VAT	VAT-reduction for cutural goods and books	1 815	red
5	§ 10 StromStG	Energy Tax	Additional ecotax reduction for energy-intensive production	1 700	yellow
6	§ 53 EnergieStG	Energy Tax	Tax subsidy for combined production of heat and power	1 300	yellow
7	§ 35a Abs. 1 Nr. 1 EStG; §	PIT	Income tax credit for labor-intensive services	1 075	yellow
8	§ 20 Abs. 4 EStG	PIT	Threshold for minimum personal saving	1 059	green
9	§ 12 Abs. 2 Nr. 10 UStG	VAT	VAT-reduction for public transport	750	yellow
10	§ 50 EnergieStG	Energy Tax	Biofuel support	670	yellow
11	§ 10a EStG	PIT	Tax credit of private retirement arrangements	560	green
12	§ 5a EStG	CIT	Simplified taxation for seafreight companies	500	green
13	§ 54 EnergieStG	Energy Tax	Ecotax reduction for agriculture	440	yellow
14	§ 44 EnergieStG	Energy Tax	Tax exemption for fossil fuels used for non-energetic purposes.	400	yellow
15	§ 27 Abs. 2 EnergieStG	Energy Tax	Tax emption for kerosene	395	red
16	§ 12 Abs. 2 Nr. 6 UStG	VAT	VAT-reduction for dental technicians' services	380	green
17	§ 3 Nr. 70 EStG	CIT	REIT-Improvement-Tool	325	red
18	§ 2 InvZulG 2007	CIT	Investment incentive for East Germany	285	green
19	§ 13 des 5. VermBG	PIT	Employees' savings tax credit	260	red
20	§ 55 EnergieStG	Energy Tax	Tax exemption for fossil fuels used for non-energetic purposes.	170	yellow
<b>R Red light</b>				4 795	5
<b>Y Yellow light</b>				10 540	10
<b>G Green light</b>				2 784	5

form unprecedented, evaluation endeavor produces some conclusions that go beyond the actual event. They concern the question of what we have learned in the project that could be important for the development of the (German) tax and subsidy policies.

The transparency of tax expenditures proves an overarching issue. Not only for evaluation purposes, but for the sake of modern and well-balanced tax policy and transfer policy, the transparency of tax breaks and reductions must be improved dramatically. The integration of tax expenditures into regular budgets, alongside direct expenditure — as recommend as best practice by the OECD (2004) — would probably produce a decisive shift towards more transparency. An additional key requirement for more transparency in Germany is openness and willingness to cooperate of all authorities concerned, especially in the states. The decentralized (also by other means controversial) tax administration of direct taxes must not serve to protect the lack of transparency in tax reliefs.

Tax expenditures in common taxes with revenue sharing between central government, states, and partly even local government raise special issues. Of the 20 tax expenditures evaluated, 12 are established in common taxes (all of them, except the energy tax reliefs). Tax expenditures in common taxes can be characterized as hidden joint tasks. These deserve further discussion on federalism reform because joint tasks pose severe governance problems. On the side of direct budget expenditure, joint tasks already have been strongly reduced.

Finally, the question arises as to whether and how the evaluation process should be continued from the evaluator's perspective. As we evaluate the project as a scientific success, an extension of evaluations to "smaller" tax expenditures seems advisable. Here, a full stocktaking of all actual tax relief — not only of those listed in the official subsidy reports — would offer the best starting point. The quick review of all British tax reliefs conducted by the new Office of Tax Simplification (2011) offers a good example when it independently identified 1042 tax reliefs in the UK tax system.

With regard to the evaluation reports submitted, monitoring, following the example of Swiss subsidy reports,<sup>7</sup> would be desirable. There it is fully tracked, whether and to what extent the recommendations of the evaluation have been implemented. In Switzerland, the law has a different character, since the subsidy report including recommendations for change is formally adopted by the Federal

Government and acknowledged the Parliament. Nevertheless, a different form of monitoring for Germany would be desirable to ensure that the evaluation efforts have an enduring effect.

Yet for this to happen, other and more fundamental prerequisites have to be met. The first is a policy background with a positive attitude to modern instruments of governance, like evaluation, and with an equally critical view of "traditional" stakeholder policy by the means of subsidization. This does not go without saying; rather, it is the exemption to the rule. This exemption ruled when the evaluation of the twenty biggest tax expenditures was commissioned in 2007. When it was finished in 2009, "the rule ruled again." The only thing that happened to the twenty tax expenditures evaluated was a massive expansion of the biggest measure which subsidizes craftsmen's services — a measure which received a deep yellow, almost orange traffic light. Thus, more than two years after the original evaluation reports were presented to the public, one might insert an additional requirement to Burman's claim cited in the in the introduction of this paper: periodically evaluating tax expenditures and drawing conclusions from this endeavor is a necessary (still not sufficient) requirement for good government.

But of course, evaluators neither make political decisions, nor do they take political responsibility for them. Evaluators take the responsibility for paving the ground for good decisions based on impartial and scientifically valid information.

## Notes

- <sup>1</sup> At least those which fall under the definition of tax subsidies used in the biannual Federal Subsidy Report. This definition is somewhat narrower than the usual understanding of tax expenditures. See FiFo/Copenhagen Economics/ZEW (2009): p. 46 f.
- <sup>2</sup> Officially, tax-free allowances for Sundays, holidays, and night work are meant as pure income transfers the persons affected.
- <sup>3</sup> See Thöne (2003), and BMF (2003), Neunzehnter Subventionsbericht, p. 13.
- <sup>4</sup> Officially though, this is not the case. Germans pride themselves of their very systematic approaches to all tax matters. Indeed, many scholars and the Federal Constitutional Court emphasize the overarching importance of the ability-to-pay principle for almost all taxes, not only for income taxes. Even when, with the tax deferral of pension claims, an obvious expenditure-tax element was introduced, a special rule called the "correspondence principle" was found to explain why this exemption need not be considered as a deviation

from SHS. But apart from this slight pedantry, real-life German tax law displays as much “pragmatic muddling through” as any other countries’ taxes do.

- <sup>5</sup> See Presidential Budget FY 2007, Appendix to Analytical Perspectives.
- <sup>6</sup> See Nationaler Normenkontrollrat (2008); Bundesregierung (2008).
- <sup>7</sup> Starting with the 3rd Subsidy Report of 2008, tax expenditures are covered, too (see Bundesrat, 2008). Still, due to a lack of empirical knowledge, tax subsidies are not reviewed intensively, so far.

## References

Bundesministerium der Finanzen

19. Subventionsbericht der Bundesregierung. Berlin, 2003.
21. Subventionsbericht der Bundesregierung. Berlin, 2007.
23. Subventionsbericht der Bundesregierung. Berlin, 2011.

Bundesregierung. Zeit für das Wesentliche: Bürokratieabbau schafft Freiräume. Bericht der Bundesregierung 2008 zur Anwendung des Standardkosten-Modells und zum Stand des Bürokratieabbaus. Berlin, 2008

- Burman, Leonard E. Is the Tax Expenditure Concept Still Relevant? *National Tax Journal* LVI(3) (2003): 613-627.
- Dahlby, Bev. *The Marginal Cost of Public Funds*. Cambridge, MA: MIT Press, 2008.
- Nationaler Normenkontrollrat, Leitfaden für die Ex-ante-Abschätzung der Bürokratiekosten nach dem Standardkosten-Modell. Berlin, 2008.
- Office of Management and Budget (OMB). Budget of the United States Government FY 2007– Analytical Perspectives. Washington DC, 2007.
- Office of Tax Simplification. Review of tax reliefs. Final report. London, March 2011.
- The Organisation for Economic Co-operation and Development. Best Practice Guidelines – Off Budget and Tax Expenditures. GOV/PGC/SBO(2004)6, Paris 2004. Schweizerischer Bundesrat, Subventionsbericht 2008 des Schweizerischen Bundesrats, 30. Mai 2008, Bern. BBl. 6229-6323.
- Thöne, Michael. *Subventionskontrolle. Ziele – Methoden – internationale Erfahrungen*. Berlin, edition sigma, 2003.
- Thöne, Michael, ed. Evaluierung von Steuervergünstigungen. Bd. 1-3, Köln, 2009.