Commentary on Burman, Toder, Berger, and Rohaly: “Economic and Distributional Effects of Tax Expenditure Limits”

Louis Kaplow*

April 19, 2016

*Harvard University and NBER. I am grateful to Alan Auerbach for comments, Jonathan Westreich for research assistance, and the John M. Olin Center for Law, Economics, and Business at Harvard University for financial support. The analysis here is applicable to many of the policies examined elsewhere in this conference volume – specifically, Williams, Dynarksi/Scott-Clayton, Hoynes/Rothstein, and Friedman – and casts a different light on a number of their points relating to distribution, distortion, and the purportedly wasteful use of government revenue. Disclaimer: I occasionally consult on antitrust cases, and my spouse is in the legal department of a financial services firm.
The tax expenditure limitation proposals that Burman, Toder, Berger, and Rohaly (BTBR) examine are advanced not only because they reduce distortions induced by the relative preference for some types of expenditures over others, but also for other reasons: efficiency gains from enabling lower marginal tax rates on account of the broader tax base; improvements in the distribution of income on account of overturning the “upside down” effect of deductions and exclusions; and enhanced revenue.¹ This commentary aims to cast a different light on BTBR and the proposals they analyze by application of modern extensions to the optimal taxation framework that employ a distribution-neutral methodology for assessing tax reforms. (For a more extensive assessment of how this framework applies to tax expenditure limitation proposals, see Kaplow (forthcoming).)

Section 1 states this methodology, which begins with the Mirrlees (1971) optimal taxation model as augmented by Atkinson and Stiglitz (1976) to incorporate commodity taxation. Tax expenditures provide subsidies to different commodities and hence are encompassed by this model. Next is a sketch of the distribution-neutral framework (e.g., Kaplow 2004, 2008).

Section 2 uses this methodology to disentangle the distinctive efficiency consequences due to the reform of tax expenditures from the efficiency, distribution, and revenue consequences of changing overall features of the tax system – changes that can be made independently of whether tax expenditures are reformed. The analysis provides a clearer positive and normative understanding of tax expenditure reform and reveals how a number of commonly held views are misleading. Prominent analysts argue that limiting tax expenditures enables one to enjoy a free lunch: raising revenue without raising tax rates, reducing the distortion due to high marginal tax rates, or generating enhanced progressivity without raising the distortion ordinarily associated

¹ For proposals and related analyses, see also Feldstein (2015), Schizer (2015), and Toder, Rosenberg, and Eng (2013).
with redistributive taxation.\textsuperscript{2} Unfortunately, these lunches are not free and, as often served, are not even available at a discount.

1. Framework

1.1 Optimal Income and Commodity Taxation

Mirrlees (1971) launched the modern study of optimal income taxation (which includes transfers to the poor), and his approach has provided the backbone for much subsequent work in public economics that seeks to ground policy analysis in first principles. Atkinson and Stiglitz’s (1976) important extension combines the Mirrlees optimal income tax problem with commodity taxation. Note that differential commodity taxation includes the case of relative subsidies (the commodity taxes may be negative, and it includes systems such as a VAT that provides preferential rates or exclusions). Moreover, the notion of “commodities” is understood broadly as encompassing all forms of expenditures and hence, combined with the first point, readily incorporates tax expenditures. Indeed, this application does not extend the commodity tax framework but merely recognizes that tax expenditures are in fact a means of differential commodity taxation.

Atkinson and Stiglitz’s central result was that, when individuals’ (taken to be common) utility function is weakly separable in labor, then it is optimal to employ uniform commodity taxation. The basic intuition comes from principles of second best analysis. Here, the preexisting distortion is of the labor-leisure choice. Hence, with weak separability – meaning that relative changes in an individual’s consumption bundle do not affect labor supply directly –

\textsuperscript{2} Consider BTBR: Tax expenditures “require tax rates to be higher than they would otherwise be, which exacerbates the efficiency cost of taxation.” (34) “Tax expenditure limitations combined with rate reductions can be designed that make the tax system more progressive [and] reduce marginal tax rates on work and saving . . .” (37) And Feldstein (2015): “Limiting tax expenditures would raise revenue without increasing marginal tax rates.” (1) “The two percent cap would also lower the marginal tax rate of all the affected taxpayers.” (5)
there is nothing to be gained by any distortion of consumption allocations. All that remains is the simple efficiency cost of consumption distortion from any differential commodity taxation.

1.2 Distribution-Neutral Approach

There is a longstanding tradition in public economics of employing a revenue-neutral approach when assessing reform proposals (other than those aimed at raising or reducing revenue). In a similar spirit, some work – and much of my own writing over the past two decades – has advanced a complementary, distribution-neutral approach. See, for example, Kaplow (1996, 2004, 2008, 2012). In a nutshell, the distribution-neutral approach combines a reform proposal with an adjustment to the income tax schedule that is designed to be distributively offsetting at all income levels.

When two different studies employ this methodology to evaluate the same reform, any differences in conclusions will correspondingly reflect differences in the assessment of the distinctive features of the core reform, rather than different choices in how to balance the budget. As BTBR amply demonstrate, distributive and distortionary consequences of a given tax expenditure limitation proposal depend on how the revenue is rebated. (Likewise for environmental reforms and many other policies considered in this conference volume.) Even with revenue-neutrality, there are an infinite variety of ways to balance the budget, with all manner of consequences for distribution and distortion (although, as will be explained, the two are tightly related). Hence, the gains from disentangling the analysis of particular reforms from broader fiscal issues are immense.

Consider some contemplated reform. As stated, distribution-neutral implementation involves adjustment of the income tax schedule so as to offset the reform’s distributive effects on utility for all levels of before-tax income. This income tax adjustment has two components. First, it washes out any effects on taxes paid (or transfers received) as a mechanical consequence
of the reform. For a tax expenditure limitation, one simply computes how much taxes paid rise at each level of income and adjusts (reduces) income tax rates accordingly. Second, because the income tax adjustment is, in principle, designed to hold individuals’ utility constant, the full income tax adjustment must also offset any other effects of the core reform on utility. Suppose, for example, that a tax expenditure proposal reduces only inefficient subsidies on certain forms of consumption. In that case, the utility consequence of reducing consumption distortion is to raise utility. Hence, the complete adjustment would involve somewhat higher income tax rates than those described in step one.

Having described what the distribution-neutral experiment is, let us now examine its consequences. Regarding distribution, it is obvious that there is no effect, by construction.

Next, consider labor supply, which is usually taken to be a first-order factor for many policies, including significant reforms of tax expenditures. Under a distribution-neutral approach, labor supply effects recede. More precisely, if one assumes that labor effort is weakly separable in individuals’ utility functions, as in Atkinson and Stiglitz’s (1976) demonstration of when uniform commodity taxation is optimal, then labor supply effects are nil. Because distribution-neutral implementation holds utility constant for each level of earnings and hence for every level of labor effort that individuals might choose, the same level of labor effort will continue to maximize utility. See Kaplow (2004, 2006, 2008).

Finally, what is the effect of the reform on revenue? The answer is immediate from examination of the two components of the distribution-offsetting adjustment to the income tax schedule. The first component is a wash: at each level of income, we adjust income tax rates to offset the mechanical effects of the core reform on taxes paid. For proposals that reduce the availability of tax expenditures, revenues rise as a direct consequence, so income tax rates are accordingly reduced to, in essence, rebate the proceeds. The second component of the income
tax adjustment absorbs the utility consequences of the reform (aside from the aforementioned mechanical effect of the change in tax payments). Because tax expenditure limitations reduce consumption distortions, they raise utility, so the second component of the income tax adjustment involves raising the income tax schedule to absorb this utility gain at each level of income. Combining these two components, we have an unambiguous increase in revenue.

In sum, the net revenue effect of the distribution-neutral reform package is the only effect of the overall reform, and it corresponds to the efficiency gain or loss produced. How this revenue change is dealt with is, strictly speaking, outside the distribution-neutral experiment. For efficient reforms, which raise revenue, one might choose to distribute the surplus pro rata, generating a Pareto improvement.

2. Application

2.1 Different Types of Tax Expenditures

At the risk of oversimplification, it is helpful to distinguish three categories of purported tax expenditures. First are those income tax provisions that are best understood not as true tax expenditures but rather as refinements to income measurement. Classification is controversial, most notably because of disagreement about the proper normative baseline. Such questions are set to the side here.³

Second – and the focus here – are provisions that provide inefficient subsidies to certain forms of consumption. Application of section 1’s framework to reforms that eliminate such tax expenditures is straightforward. As explained, the distribution-neutral approach involves an adjustment to the income tax schedule that has two components. First, tax rates are reduced sufficiently at each income level to leave taxpayers with as much disposable income as before

³ Under a more rigorous optimal income tax analysis, such taxonomic questions have no direct role; the optimal treatment of any item is whatever comes out of the mechanism design exercise. Relatedly, because BTBR largely exclude capital taxation from their analysis, the proper way to incorporate that subject is less important for this essay. For further exploration, see Kaplow (2008, ch. 9).
their tax expenditures were limited. Second, tax rates are increased sufficiently at each income level to absorb the utility gain from eliminating the distortion to individuals’ expenditure decisions. Distribution-neutrality holds by construction. With regard to revenue, the first component alone entails revenue neutrality, and the second component results in a revenue gain that equals (in dollars) individuals’ aggregate utility gains from the reduction in expenditure distortion.

A third category of tax expenditures includes those that correct otherwise distorted behavior. For example, the charitable contribution deduction may be seen as a Pigouvian subsidy on a class of expenditures that generates positive externalities. Analysis of this case involves a modest extension of section 1’s basic framework. See Kaplow (2004, 2012). The core idea is that the subsidies should be moved toward the optimal Pigouvian correction (which is zero when there is no externality). If the charitable deduction is too generous in this regard, then limiting it would produce a gain, whereas if the deduction provided too little benefit, then further limitation would generate a loss.4

2.2 Non-Distribution-Neutral Reforms

To many, one of the most apparent and disturbing features of tax expenditures is what Surrey (1973) referred to as their “upside-down” effect. Because exclusions and deductions have a value determined by individuals’ marginal tax rates (MTRs) – and for other reasons – many tax expenditures are significantly more favorable to higher-income taxpayers. Accordingly, many favor tax expenditure limitations in order to augment progressivity.

As emphasized by Griffith (1989), however, this perspective is overly simplistic. After all, the existing regime did not take some tax rate schedule from Plato or Pluto, set it in stone,

---

4 As an aside, for atmospheric externalities, the optimal Pigouvian subsidy, when administered through the income tax, would take the form of a refundable credit (with no limit). One can also extend the analysis to the correction of so-called internalities. For further discussion, see Schizer (2015).
and then have someone else (Congress) superimpose tax expenditures. Rather, a single political process generated all the features of the existing regime. Moreover, this regime is reformed from time to time: sometimes the tax rate schedule, sometimes various of the tax expenditures, and sometimes a combination. An example of particular note is the Tax Reform Act of 1986, which (speaking roughly) broadened the base by reducing tax expenditures and simultaneously lowered rates, and in a manner that linked the two together, specifically, to achieve distribution neutrality (as well as revenue neutrality). In such a case, there is by construction no relationship between the level of tax expenditures and progressivity.

Suppose, however, that one wishes to understand and assess non-distribution-neutral reform packages, perhaps of the sort that some reformers envision under which tax expenditures would be significantly limited but the tax rate adjustments would result in an overall greater degree of redistribution. One can augment the distribution-neutral methodology using a simple two-step decomposition. See Kaplow (1996, 2004, 2008).

1. Distribution-neutral implementation: Combine the tax expenditure limitation component with a distribution-neutral income tax schedule adjustment of the sort examined throughout.

Then, an instant after this is to be enacted, and with the same effective date, do the following:

2. Pure redistribution: Implement an adjustment to the income tax schedule that moves from the distribution-neutral schedule to the actual schedule in the overall proposal under consideration.

To analyze any non-distribution-neutral reform, one can simply combine the analysis of these two steps.

Analysis of the first step is straightforward from the previous discussion. It is distribution neutral by construction. It does not affect labor supply. Its only consequence is the
pure efficiency effect of the tax expenditure limitation with regard to reducing individuals’
consumption distortions.

Analysis of the second step is readily understood once one appreciates what it is: a purely
redistributive change to an income tax system. Accordingly, the correct analysis is generic: it is
that from the standard optimal income tax problem. We know that if step two involves an
increase in redistribution, for example, we have whatever social gains are deemed to be
associated with that change and the social cost of increased distortion. A related point, to be
elaborated in subsection 2.4, is that if effective MTRs fall and hence distortion falls, then we
know that we have less redistribution.

Note further that, being generic, this second step could be achieved in a variety of ways,
including without regard to tax expenditure limitations. That is, step two – some particular
reform to the income tax schedule that involves purely a change in redistribution – could be
implemented by itself. Or packaged with a change in the military budget, spending on highways,
or the regulation of greenhouse gas emissions.

The distribution-neutral framework, even when one is not examining an actually
distribution-neutral reform package, facilitates specialization and clarifies communication. If a
single study aims to analyze the total effects of a non-distribution-neutral package – whether of
tax expenditure limitations, road construction, or a gasoline tax – it must take on all the burdens
of those who analyze optimal income taxation: making assumptions about elasticities, choosing a
social welfare function, and so forth. If it instead analyzes the distribution-neutral version, it can
focus on the distinctive features of the subject at hand. (Any distributive effects of the overall
package might be displayed, but left for others to analyze.)

Closely related, it can be very difficult to compare two studies of the same subject if
distribution-neutral implementation is not employed. As BTBR explain, there are many (indeed
an infinite number of) ways to achieve revenue neutrality. And even if the two studies make the same assumption about redistribution (they are analyzing the same overall package), it may be difficult to disentangle whether, say, the first study’s more favorable bottom line is due to its more positive assessment of the distinctive features of the reform or perhaps a more negative assessment of those but a more positive view of the resulting change in redistribution. By contrast, the use of distribution-neutral implementation provides a common metric that enables apples-to-apples comparisons.

### 2.3 The Relationship between Tax Expenditure Limitations and MTRs

Whereas the previous subsection elaborated on distribution, this one will focus on the distortion side of the familiar distribution-distortion tradeoff. The two-step decomposition makes clear that changes in both distribution and labor supply distortion will be located in the second, purely redistributive step of a reform. Unfortunately, this linkage is often forgotten. Some proponents of tax expenditure limitations believe that they can have their cake and eat it too – in economists’ parlance, a free lunch (rather than dessert). We should not be surprised to discover that such is not the case.

Consider a linear income tax with marginal tax rate $t$. Furthermore, there are deductions or exclusions for tax expenditures that in aggregate are the fraction $\alpha$ of individuals’ gross labor earnings $y$. Tax due is $t(1-\alpha)y$.

Now, examine a global tax expenditure limitation that allows only the fraction $\theta$ of tax expenditures to remain exempt. If the revenue is all used to reduce $t$, the new statutory tax rate becomes $\frac{1-\alpha}{1-\theta\alpha}$. However, simple algebra shows that the tax due after the reform remains $t(1-\alpha)y$. Therefore, the following are true: (1) Individuals at all income levels pay the same amount of tax that they did before; this implies both revenue- and distribution-neutrality. (2) The
statutory MTR falls from \( t \) to \( \frac{1-\alpha}{1-\theta \alpha} t \), which is a lower rate because we are assuming that \( \theta < 1 \).

(3) The effective MTR stays the same, at \( t(1 - \alpha) \). This latter point is key. Note that the statutory MTR of \( t \) was not the effective MTR to begin with. After the reform, the effective MTR is unchanged: the new statutory MTR is \( \frac{1-\alpha}{1-\theta \alpha} t \), which is applied to the fraction of income \( 1 - \theta \alpha \), giving the same effective MTR of \( (1 - \alpha)t \).

Suggestions that broadening the tax base through tax expenditure reform enables lower MTRs are either misleading or incorrect. If interpreted as reductions of the statutory MTR, they are accurate but convey the misleading impression that the effective MTR and hence the distortion of the labor-leisure margin are lower. If interpreted as reductions in the effective MTR, they are incorrect.

2.4 Progressivity and Distortion

A recurring theme of this essay is that neither black magic nor tax expenditure limitations enable us to escape the distribution-distortion tradeoff inherent in redistributive taxation. Distribution-neutral implementation leaves both distribution and labor supply distortion unchanged, features shared by the example in the preceding subsection. The two-step decomposition makes clear how deviations from distribution-neutrality can be isolated, leaving (as step two) a purely redistributive change to the tax system. And that is where our familiar distribution-distortion tradeoff resides. Because some proponents of tax expenditure limitations present a rosier picture of the possibilities – suggesting that one can enhance progressivity, for example, without raising or even while lowering the core distortion of the income tax – some elaboration is useful.

For concreteness, let us employ a standard definition of progressivity (PROG): rising average tax rates (ATRs). That is, \( \text{PROG}(y) = \frac{d\text{ATR}(y)}{dy} \). Taking that simple derivative, we
have \( \text{PROG}(y) = \frac{\text{MTR}(y) \times \text{ATR}(y)}{y} \). In the present discussion, \( \text{MTR}(y) \) and \( \text{ATR}(y) \) are both taken to be *effective* rates, in the spirit of the preceding subsection’s analysis.

This expression teaches a straightforward lesson: the only way to increase the level of progressivity at a given income level – which is to say, to increase \( \text{ATR}(y) \) faster than before – is to push \( \text{MTR}(y) \) higher than before. Hence, the suggestion that we can increase progressivity while maintaining or reducing MTRs has to be mistaken. As subsection 2.3 just explained, a significant part of the problem may involve confusing statutory and effective MTRs.

The only remaining wrinkle is how the efficiency gain from eliminating consumption-distorting tax expenditures might be distributed to the population – which is taken to be a choice that lies outside the distribution-neutral experiment. Note that if it was rebated pro rata, effective MTRs are unchanged but, actually, PROG rises because, like the lump-sum component of any tax schedule, an additional dollar is a greater percentage of income the lower is one’s income. Hence, the core efficiency gain from tax expenditure reform does enable a free lunch. But that, indeed, is the entire point of basic efficiency analysis.

### 2.5 Raising Revenue

Raising revenue is another often-advanced goal of tax expenditure limitation proposals. This too is best understood by applying the distribution-neutral framework. With distribution-neutral implementation, as we were just reminded, revenue effects arise entirely on account of correcting distortions caused by tax expenditures. Most who advance tax expenditure limitations on revenue-raising grounds, however, are not referring to this aspect, but instead have in mind retaining some or all of the revenue mechanically raised by the tax expenditure limitations in order to fund programs or reduce the deficit.

To analyze this further, consider the following variant of the two-step decomposition from subsection 2.2:
1. Distribution-neutral implementation: Combine the tax expenditure limitation component with a distribution-neutral income tax schedule adjustment of the sort examined throughout. Then, an instant after this is to be enacted, and with the same effective date, do the following:

2. Pure revenue-raising: Implement an adjustment to the income tax schedule that moves from the distribution-neutral schedule to the actual schedule in the overall proposal under consideration.

The only difference from the description of the original two-step decomposition is that the label “pure redistribution” is now changed to “pure revenue-raising.” Again, the second step – here, a pure increase in (effective) tax rates to fund whatever – can be analyzed generically. That is, the analysis would be essentially the same if policy-makers implemented step two without regard to whether it had anything to do with a tax expenditure limitation proposal. Higher taxes may be a good thing. Or they may not. But the answer does not depend on whether taxes are raised in isolation, as part of a tax expenditure limitation, as part of the use of proceeds from a carbon tax, or in some other manner.

It might be thought that revenue-raising is less distortionary, all else equal, after a tax expenditure limitation is implemented because one starts with lower MTRs. However, as explained in subsection 2.3, this view confuses statutory and effective MTRs. If one enacted the tax expenditure limitation and did not use the proceeds to reduce tax rates, as in the prior illustration, then effective MTRs would rise rather than be constant. The only way to keep effective MTRs the same is to rebate the proceeds in a distribution-neutral fashion, leaving no revenue.

There is, however, an important respect in which revenue-raising (and redistribution, if one so chooses) may become more economically efficient as a consequence of tax expenditure
limitations. Even though the benchmark effective MTR on labor income is unchanged – no free lunch there – it is the case that when a tax system has a broader base, it may well be less distortionary to raise MTRs. In a pure income tax regime, raising the MTR distorts labor supply and nothing more. In an income tax littered with distortionary tax expenditures – specifically ones that are in the form of exclusions and deductions and hence a function of statutory MTRs – raising statutory MTRs causes additional consumption distortions as well as labor supply distortion. For further analysis, see Kopczuk (2005) and Kopczuk and Slemrod (2002).

2.6 Distribution Neutrality as a Descriptive Benchmark

BTBR do an impressive job not only in modeling the effects of tax expenditures on revenue and tax burdens at various levels of income but also in displaying the results of their analysis. The greatest challenge they face in this respect is that different formulations of tax expenditure limitations and different ways of rebating the revenue raised from such limitations have important effects on the resulting distributive outcomes. Moreover, we must intersect each limitation proposal with each method of using the revenue to cut tax rates (and there is an infinity of ways to do the latter).

A distribution-neutral benchmark can help. One simple way to display a tax expenditure limitation’s consequences is to compare statutory MTRs under it to those under distribution-neutral implementation. The pattern of differences would show the distributive effect of the limitation proposal, setting aside how the proceeds are used. Alternatively, one might graph total taxes paid as a function of income for each of the three tax expenditure limitation proposals that BTBR examine and for distribution-neutral implementation (which is equivalent to the status quo). This single graph would clearly depict the differing distributive effects. A picture is worth (at least) a thousand words, and they can be absorbed more readily than if the thousand words (numbers) appear in one or more dense tables.
One can proceed similarly in order to compare the effect of various ways of reducing tax rates in order to rebate the revenues raised by a tax expenditure limitation proposal. That is, one can present taxes paid as a function of income under distribution-neutral implementation and compare it to the schedule under each alternative. Likewise, one could present the distribution-neutral schedule and the schedules for each of the posited alternative means of returning the revenue on a single graph, making clear all the differences.

BTBR also consider how various packages might affect incentives, and here I will focus on labor effort. The aforementioned graphs provide this information as well. We know that, in the benchmark case, the distribution-neutral version leaves labor effort unchanged. Hence, wherever the slope of the tax schedule under an alternative scheme is higher (lower) than under distribution-neutral implementation, work incentives are reduced (increased). Or, more directly, by graphing the effective MTRs under the distribution-neutral version and any other, one can immediately see how labor incentives change.

Note that each of these comparisons between tax schedules under various combinations of expenditure limitation and tax rate reduction to those under distribution-neutral implementation of the corresponding tax expenditure limitation are simply a depiction of the two-step decomposition introduced previously. Step one is the distribution-neutral curve. Each of the other curves combines steps one and two. Hence, the difference between the distribution-neutral curve and another curve will depict step two – the purely redistributive component – in isolation.

---

5 The text refers to the substitution effect because it is MTRs that are being compared. To capture the uncompensated change, one would also have to integrate the effects, or examine the schedules for total taxes paid, to extract the income change and then determine the income effect.
3. Conclusion

The distribution-neutral framework for policy analysis that is developed in a body of work over the last two decades is an extremely powerful tool for the analysis of tax expenditure limitations. And the two-step decomposition allows one to isolate the distinctive effect of tax expenditure reform – the reduction in distortions of expenditure choices – from effects on distribution, labor supply distortion, and revenue. The analysis shows how a number of commonly advanced beliefs about the benefits of tax expenditure limitation proposals – including the view that there exists a free lunch wherein one can reduce MTRs and enhance revenue or redistribution simultaneously – are largely illusory.

Although outside the scope of this essay, a distribution-neutral approach may also have some bearing on the construction of politically feasible reform packages. The Tax Reform Act of 1986 ostensibly broadened the tax base via the reduction of tax expenditures in a manner that was not only revenue-neutral but also distribution-neutral. As explained, distribution-neutral packages result in Pareto improvements (more realistically, gains on average to every slice of the income distribution) when the underlying reforms are efficient. Such a package may be politically feasible.
References


