WP/13/57

# IMF Working Paper

# Targeting, Cascading, and Indirect Tax Design

*Michael Keen*

INTERNATIONAL MONETARY FUND

 WP/13/57

# IMF Working Paper

Fiscal Affairs Department

# Targeting, Cascading, and Indirect Tax Design*

**Prepared by Michael Keen**

Authorized for distribution by Michael Keen

February 2013



## Abstract

This paper addresses two fundamental issues in indirect tax design. It first revisits the case for reduced rates on items especially important to the poor, establishing conditions under which even very crudely targeted spending measures better serve their interests. It then explores the welfare costs from cascading taxes, showing that these may actually be lower the wider the set of inputs that are taxed but, more to the point—and contrary to the common notion that "a low rate on a broad base" is always good tax policy—may plausibly be large even at a low nominal tax rate and with few stages of production.

JEL Classification Numbers: H21, H23

Keywords: Indirect taxation, Value Added Tax, Targeting, Cascading

Author's E-Mail Address: mkeen@imf.org

* This is a revised version of the Third Dr. Raja J. Chelliah Memorial Lecture, "Some Current VAT Issues," delivered in New Delhi on February 9, 2012. I am grateful to Govinda Rao for the kind invitation to deliver this lecture; to him, Kavita Rao, and others of the National Institute of Public Finance and Policy for their generous hospitality and comments on this work; and to David Coady for many helpful suggestions. The views expressed here remain mine alone, and should not be attributed to the IMF, its Executive Board, or its management.

Contents Page

# I. INTRODUCTION

*In short, the Committee recommended… moving towards a full-fledged VAT.*

Chelliah Report,[1] p.5

In a long and distinguished career,[2] Raja Chelliah left many marks on the profession of public finance, in India and beyond. Not the least of these was in his advocacy for India of the value-added tax (VAT). The 'Chelliah Report' remains not only a classic statement of the specific case for the adoption of a VAT in India but also one of the most clearly and elegantly argued assessments of the VAT's generic strengths and weaknesses. Twenty years later, the VAT in India—or rather, as it is known, the General Sales Tax (GST)—remains a work in progress and the topic of considerable controversy. So it seems both appropriate and timely that this lecture focus on the design of and case for (or against) the VAT. The aim is not to address the wider range of challenges that VAT reform in India faces,[3] but to focus on two general issues that arise wherever the VAT is discussed (and, as will be seen, even more widely than that). Both, needless to say, were well known to Raja Chelliah. The purpose here is simply to take the understanding of each a little further.

The first issue, considered in Section II, is that of whether a reduced rate of VAT on some commodities that account for a particularly large part of the spending of the poor—'food' is the classic example[4]—can make good sense as a way of addressing equity concerns. In India, as elsewhere, this has been at the forefront of public debate: the Empowered Committee of State Finance Ministers (2009), for instance, rather hedged its bets by suggesting a standard rate of 20 percent (state and federal VATs combined) and reduced rate of 12 percent on necessities—but converging to 16 percent within three years. Analysts often point out that most of the benefit of reduced indirect tax rates actually accrues to the better-off, making this a very poorly targeted way of pursuing equity objectives. But that cannot be the end of the argument. In less advanced economies, it is not obvious that there are any better ways to protect the poor available to the government. So the aim here is to take the argument one step further, by asking: How well targeted do public spending measures have to be for the poor to be best served not by taxing at a particularly low or zero rate those commodities that account for an especially large part of their budget, but by taxing them and using the proceeds to increase that public spending?

---

[1] Chelliah et al (1992).

[2] Including, it is a pleasure to mention, a stint in the Fiscal Affairs Department of the IMF.

[3] A thorough and persuasive account is in Cnossen (2012); Rao (2011) provides a lively perspective on the central issues.

[4] As Cnossen (2012, p. 9) puts it: "The most contentious issue under any VAT is the treatment of food."

The second issue, taken up in Section III, is that of 'cascading': the levying of tax on items that have already been taxed. The term is, effectively, used by public finance theorists as a term of abuse, its mere invocation tainting any proposal as inherently suspect. Certainly this has been a prominent concern in India, a primary argument for movement to a full-fledged VAT being—as it was in the Chelliah Report—precisely that this would substantially reduce the current extent of cascading. This is indeed considerable cascading under the present and remarkably complex indirect tax arrangements in India:[5] Poddar and Ahmad (2009) speculate that more than 35–40 percent of the revenue from these taxes may come from cascading elements.

But what exactly is wrong with cascading? And—even more rarely asked—how large might be the welfare losses that it implies? These are the questions addressed here. They have importance, it should be stressed, far beyond the case for a VAT over some form of turnover tax. Even the best designed VATs embody some degree of cascading in the form of 'exemptions': provisions by which the sale of some commodity—financial services[6] are a very common example—are not subject to VAT but nor is any refund given for the tax charged on the inputs used in their production, so that this input tax 'sticks'—with tax then levied on top of that tax when the exempted item is used as an input into production. So understanding cascading is also critical to understanding the imperfections of real-world VATs themselves. It can also be critical in relation to retail sales taxes, even though these should in principle be levied only on final sales to consumers: Ring (1989) finds that on average, about 40 percent of the revenue from the state-level retail sales taxes (RSTs) in the United States are collected on business purchases. And the issue is central even beyond consumption taxation, being critical, not least, to the current debate, especially heated in Europe, on the merits of taxing financial transactions: here too cascading will inevitably arise, both in the multiple transactions involving the same financial instrument and in so far as the cost of these transactions is embedded in final product prices to business users. At its most fundamental, the costliness (or otherwise) of cascading is central to the public finance mantra of "broad base, low rate": since the broadest base will involve taxing as wide a range of transactions as possible, even if tax is again levied at a greater stage, this can only be wise as unqualified advice if cascading is, in fact, not a very great concern.

---

[5] Cascading elements include the delay in crediting tax on investment spending under the CENVAT, the exclusion of major sectors (including agriculture, real estate construction, oil and gas production) from the CENVAT and the denial to them of credits on state VAT paid, and the non-creditable Central Sales Tax on inter-state trade.

[6] The difficulty arises particularly for financial services charged for in the form of a margin, since it is then unclear how to allocate the value added (as given, for instance, by the spread between borrowing and lending rates) between the two sides of the transaction (lender and borrower), as is needed for the invoice-credit form of VAT to ensure that tax is appropriately credited to business users but not to final consumer. Exemption was seen as a way of ensuring that at least some revenue was collected in respect of such services, though it is now better understood how such services might in principle be more properly brought into the VAT. See, for example, Poddar and English (1997).

## II. When are the Poor Best Served by a Reduced Rate?

A perennial issue in VAT design is the question of whether the applicable rate should be the same for all goods and services—a 'uniform' rate structure—or whether it should vary across them. The most prominent question in practice is whether items that are especially important to the poor should be taxed at a differentially low (or zero) rate.[7] India, as noted at the outset, is no exception. The professional consensus, it is fair to say, has been in favor of uniformity. While such 'expert opinion', as Bird and Gendron (2007, p. 222 ) warn, "is biased…by the particular experience of the experts in question" there is in this case a substantial body of thought and, to a lesser extent, empirical evidence behind the standard view. It has, in any event, proved increasingly influential: as Table 1 shows, the proportion of VATs which were introduced with a single rate had increased markedly over time, to a point at which uniformity, on introduction at least, has become the norm.

**Table 1. VATs with a Single Rate at Time of Introduction**

| | Number of new VATS | Percentage with a single rate at introduction |
|---|---|---|
| Before 1990 | 48 | 25 |
| 1990–1999 | 75 | 71 |
| 1999–2011 | 31 | 81 |

Source: IMF data.

The many conceptual and practical arguments for and against various forms of differentiation in commodity taxation have been extensively reviewed elsewhere.[8] The discussion here will focus only on the distributional concern raised above, the question being: If some good—'food' is the archetypal example—accounts for a larger share of the expenditure of the poor than of the rich, should it on these grounds be taxed at a lower rate than the generality of commodities?

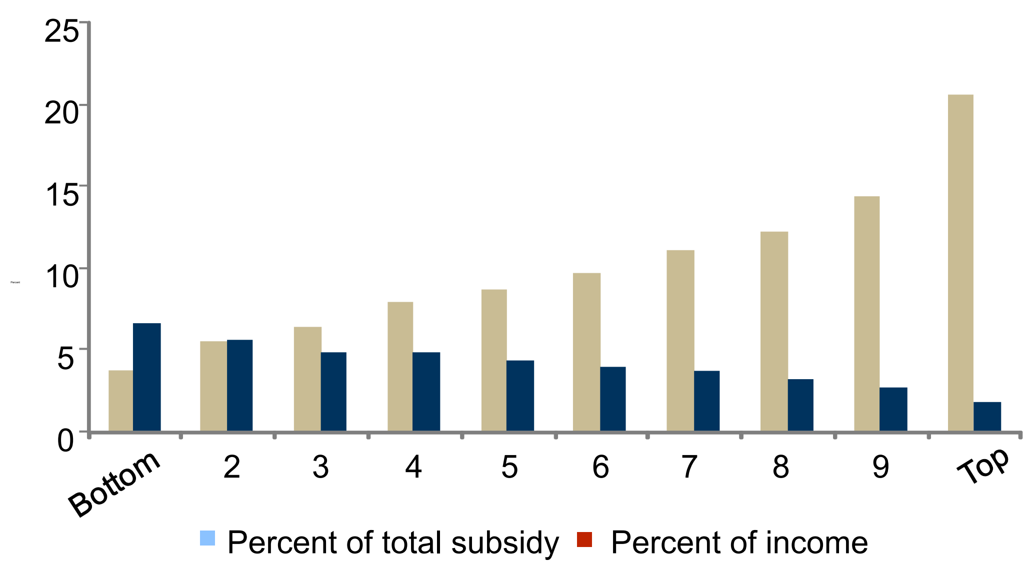
[7] More generally, of course, commodities might (and quite often are) subsidized. This possibility is encompassed by the analytics below, but will not be lingered on in the text.

[8] A relatively nontechnical account, and references, can be found for instance in Crawford, Keen, and Smith (2010).

**A common but incomplete argument**

The starting point is a simple and clear lesson from theory: the amount of redistribution that can be achieved by differentiating rates of indirect taxation will generally be quite limited. This is because variation in the share of income spent on particular goods is generally just not great enough to make this an effective way to distinguish between poor and rich.[9] Put somewhat differently, even though the poor may spend a large *proportion* of their income on (say) food, the rich are likely to spend more on food in *absolute* terms—so most of the revenue foregone by taxing it at a low or zero rate effectively accrues to the rich, not the poor. Figure 1 illustrates this for the case of Mexico, a country which has extensive zero-rating of food and other items The darker bars confirm that the implicit subsidy is indeed greatest, relative to income, for the lowest income deciles—because zero-rated commodities account for a larger part of their income. But the lighter bars show that the share of the total revenue foregone by zero rating is higher at higher levels of income—because the richer spend a large amount on zero-rated commodities. And the latter effect is strikingly large: for each $100 foregone by zero-rating, less than $5 benefits the poorest 10 percent of the population; and more than $20 benefits the top 10 percent.

**Figure 1. Distributional Impact of Zero-rating in Mexico**




Source: OECD (2007).

[9] The theoretical argument here is elaborated in Sah (1983); see also Box 7.4 of Ebrill et al (2001).

The argument against reduced VAT rates on equity grounds all too often stops here. But it is not enough to show that this is a poorly targeted way to help the poorest: if a reduced rate is the only way to protect the poor, then costly as it is, it may nevertheless be judged optimal policy. The issue is whether better instruments for helping the poor are available.

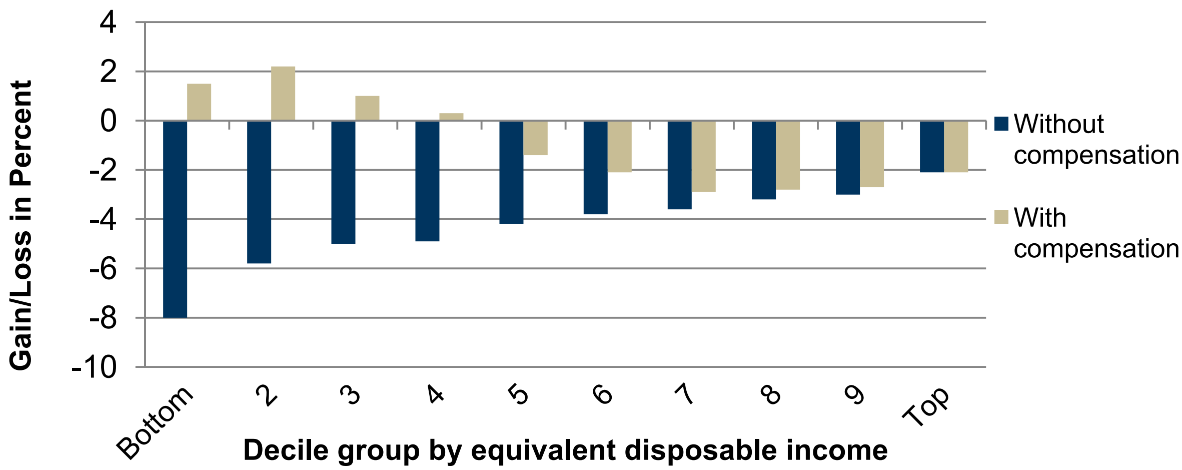
**When are spending instruments better?**

Intuition suggests, and theory confirms, that the equity argument for differential rates of commodity taxation grounds is weaker the more sophisticated are the other instruments that the government can deploy to address its distributional concerns.[10] The question then is whether the government does indeed have at its disposal a rich enough set of instruments to achieve its distributional aims by other means.

***Income-related benefits***

There is little doubt that advanced economies generally do. Take, for instance, the U.K. which, like Mexico, zero-rates a large part of consumption, including most foods. If these items were instead to be charged at the standard VAT rate, this would imply the regressive pattern of losses shown by the dark bars in Figure 2, the reason being the same as that behind

**Figure 2. Removing Zero-rating in the U.K., With and Without Compensation**




Source: Crawford, Keen, and Smith (2010). Figures based on unifying VAT at 17.5 percent in 2005–06; compensation is by a 15 percent rise in means-tested social assistance and tax credits (and costs only around half of the revenue gained from eliminating zero-rating).

---

[10] The formal hint of this is that the restrictions on preferences (assuming these to be identical across individuals) sufficient for uniformity to be optimal when the government can deploy a fully nonlinear income tax are weaker than those needed when it can only deploy a linear income tax: Weak separability between commodities and leisure (in the sense of time spent not in paid work) together with linear Engel curves in the latter case, weak separability alone in the former (Deaton (1979); Atkinson and Stiglitz (1976)).

the dark bars in Figure 1: the proportion of income spent on food and other zero-rated items decreases with the level of income, so the loss from taxing these commodities, as a proportion of income, is higher at lower income levels. But the U.K. also has a comprehensive system of means-tested (that is, income-related) transfers, and the lighter bar shows that by using part of the revenue raised by eliminating zero-rating to uprate these benefits, the losers from eliminating zero-rating can be more than compensated, with the government still enjoying a net revenue gain.[11]

***Less finely targeted public spending***

What though of less advanced countries, in which such precise means-testing is not available? How well targeted must public spending be—whether on some simple form of cash transfers or the direct provision of goods and services—if it is to provide a more effective way of helping the poor than a reduced VAT rate? The examples above (and others) suggest that the 'leakage' to the non-poor of the benefits from a reduced VAT rate that is intended to help the poor may actually be so great that even spending measures that are not very finely targeted on the poor may be a better way of helping them. But just how bad must these spending instruments be if reduced commodity tax rates are to be the best way to help the poor?

To address this as sharply as possible, suppose that the sole objective of policy is the maximin one of maximizing the welfare of the poorest individual, denoted $p$, that the only tax instruments available are commodity taxes, and that the revenue from these is the sole source of finance for a single item of public spending, in amount $G$. Suppose too that some fixed proportion $s^h$ goes to the benefit of each household $h$. In this simple framework (the details being in Appendix A),[12] starting from some arbitrary initial pattern of taxes and spending, a small increase in the tax on some commodity $k$, the proceeds from which are used to increase public spending, benefits the poorest individual $p$ if and only if

[11] While the focus here is on equity concerns, reforms of the type just described do raise important efficiency considerations, because compensating some (but not all) consumers for the effect of an increased VAT rate will generally affect incentives to provide paid work. Here, things quickly become complex, with effects arising from both the VAT reform itself and the compensating reform of the direct tax-transfer system. If the compensation through the latter were exact for everyone, overall incentives to work would be expected to be largely unchanged. (For the case in which all have identical preferences, weakly separable in the manner of the Atkinson -Stiglitz (1976) result cited earlier, Laroque (2005) and Kaplow (2006) show that they would be entirely unchanged, and that government revenue would increase—so that a subsequent direct tax reduction could ensure that the movement toward uniformity is Pareto-improving). Where means-testing is used only to protect the poorest, however, as envisaged in the U.K. example, the associated withdrawal of these additional benefits over some range will mean, for at least some, a lesser marginal incentive to undertake more paid work.

[12] The analysis here is closely related to several treatments of optimal tax and transfer design. Coady (2004), in particular, explores the poverty impact of a range of social support measures in a framework similar to the present.

$$\theta_k^p < \lambda^p S^p \left(\frac{1+\Gamma_k}{1-\Gamma_G}\right) \tag{1}$$

where $\theta_k^p \equiv C_k^p / \sum_h C_k^h$ is the proportion of all consumption of commodity $k$ accounted for by the poorest; $\lambda^p$ is the money-equivalent valuation that *p* places on an additional \$1 of public spending allocated to its benefit; $\Gamma_k$ reflects the indirect effect on revenue of an increase in the tax on commodity $k$ (the effect, that is, resulting from induced changes in consumption) and $\Gamma_G$ is the impact that an increase in public spending has, through its effect on commodity demands, on tax revenue.[13]

Understanding condition (1) is the main task in the rest of this section.[14] It may look somewhat convoluted, but in fact builds easily on the intuition above. What matters on the left is the proportion of commodity $k$ that is consumed by the poorest. This serves as a simple indicator of how well targeted is a reduced rate on $k$ is a way to benefit the poorest—as is in line with the intuition above, since this is the quantity (corresponding to the lighter bars in Figure 2, for example) that determines what proportion of the implicit subsidy associated with a reduced tax rate actually benefits them. Condition (1) then simply compares the targeting effectiveness of a reduced rate on commodity $k$ to the term on the right hand side. To understand the latter, it is useful to proceed in steps. For this, suppose, until further notice, that increasing the price of *k* has no effect on the demand for any taxed good (so that $\Gamma_k = 0$) and neither does an increase in public spending (so that also $\Gamma_G = 0$).

Given this, it helps fix ideas to start with the case in which public spending is in the form of cash transfers. In this case, \$1 of public spending allocated to its benefit means \$1 of cash received, so that $\gamma^p = 1$. Condition (1) then reduces to $\theta_k^p < S^p$: a simple comparison between indicators of the targeting effectiveness of reduced tax on *k* and the cash transfer system.

One particular structure of cash transfers that provides a useful benchmark that in which commodity tax revenue is simply returned (at least at the margin) as an equal per capita cash payment to all *H* individuals in the population. (This may seem far-fetched, but it is notable that Iran has recently done did precisely this (including by setting up bank accounts for more than sixty million people) in order to address the distributional consequences of drastically reducing petroleum subsidies). In this case $S^p = 1/H$, and the condition (1) becomes simply

[13] It is assumed throughout that $\Gamma_G < 1$, so that increased public spending does not pay for itself.

[14] This condition, it should be noted, does not speak directly to the uniformity issue. Indeed some departure from uniformity will be desirable unless, when evaluated at some uniform tax structure, the inequality in (1) becomes an equality; and it will be in the direction of a lower rate on *k* if and only if the inequality is reversed.

the requirement that $C_k^p < (\sum_h C_k^h)/H$ : increasing the rate on $k$ benefits the poorest so long as their capita consumption of $k$ is less than the per capita average consumption over the whole population—a very weak condition indeed, requiring no more than a positive income elasticity of demand for $k$.

To the extent that cash transfers can be targeted in a way more progressive than a uniform poll-subsidy, the case against especially low rate on commodities important in the budgets of the poor is of course even stronger. For many emerging economies, increasing experience with cash transfer schemes and the potential for proxy means-testing[15] is making this increasingly feasible—a point stressed in the Latin American context by Barreix, Bès, and Roca (2012), who demonstrate too the substantial reduction in poverty that could in principle be achieved by eliminating reduced rates and transferring to the poorest sufficient of the proceeds. The key issue, of course, is that of designing transfer schemes sufficiently well targeted to do this.[16]

While some form of targeting of cash transfers is likely to be possible in many countries, the precision with which this can be achieved is likely to remain limited in many lower income countries. For them, the main alternative to supporting the poor by rate differentiation is likely to be through the public provision of goods or services, such as basic health and education and infrastructure.

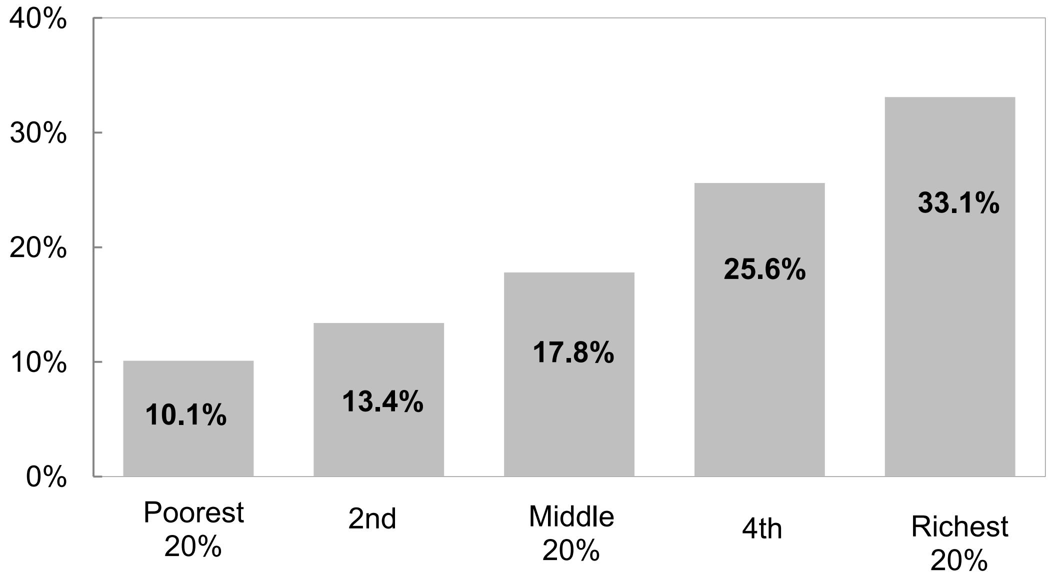
In such cases, $G$ in condition (1) must be interpreted as some such form of spending. Some guidance in understanding how this affects the case for rate differentiation is provided by empirical studies that aim to quantify distributional aspects of public spending, by allocating costs incurred to specific recipient groups; in effect, that is, estimating the $S^h$. Figure 3 shows the results, for instance, of a study of the distributional benefits of curative health care in India. The overall pattern is regressive, in the sense that the better off benefit from a disproportionately large share of the spending. In this case, only 10 percent of the spending goes to the benefit of the poorest 20 percent. To see the implications, suppose, to keep the numbers simple, that the total population size is 100, and total curative health care spending is \$100; so the poorest 20 people benefit from spending of \$10. Suppose for the moment that each of them values this at precisely its per capita cash value, 50 cents, so that again $\lambda^p = 1$. Then the right of (1) is $S^p = (20) \times 0.5/100 = 0.1$, and the uniform tax strategy is preferred if the poorest 20 percent account for less than 10 percent of all zero-rated consumption. This too is far from implausible. Indeed in one of the few empirical studies that looks at the joint effect of a uniform VAT and the public spending it might finance, Muñoz and Cho (2004)

---

[15] This uses correlates of income or need to construct indicators on which transfers might be based: see for instance, Castañeda (2005) on experience with such a scheme in Colombia.

[16] It is perhaps worth noting that while biometric cards containing only personal identifiers can be useful in ensuring that benefits are delivered to intended recipients, additional information is needed to deliver benefits more finely targeted than a poll subsidy.

**Figure 3. Distribution of Benefits of Curative Health Care**




Source: Mahal et al (2007).

find that the poorest 40 percent in Ethiopia would benefit from adoption of a uniform rate VAT, the proceeds of which are used to scale up spending on basic education and health.

All this, however, may still understate the case against differentially low rates. For the value that the poorest recipients place on public spending allocated to their benefit may well exceed the cash cost involved, so that $\gamma^p > 1$. Continuing the example of the previous paragraph, if the poorest 20 people each valued their enjoyment of the \$10 allocated to them not at its cost of 50 cents but at \$1, so that $\lambda^p = 2$ then $\lambda^p S^p = 2 \times 0.1 = 0.2$; recognizing that the valuation of public provision may differ from its cost means that increasing the tax on $k$ is the best way to support the poor under the even weaker condition that their per capita consumption is below the population-wide average.

It remains to consider the roles played by $\Gamma_k$, the indirect revenue impact of increasing the tax on $k$, and by $\Gamma_G$, the revenue impact of public spending itself.

For the former, the lesson is simply that the case for a reduced rate is strengthened to the extent that it changes consumption in such a way as to increase overall revenue, perhaps by inducing strong substitution towards other taxed goods. But the magnitude and even sign of this effect will be sensitive both to patterns of behavioral response and the pre-existing tax system. In the absence of cross price effects, for instance, $\Delta_k$ reduces to the product of the initial tax rate on $k$ (in ad valorem form) and the aggregate own-price elasticity of demand for

$k$. Starting from a position in which $k$ is untaxed $\Gamma_k$ is then zero, and the question reduces[17] to a simple comparison of $\theta_k^p$ and $\lambda^p$ as above; if the tax rate is initially 15 percent and the price elasticity of demand -0.1, the comparison is between $\theta_k^p$ and $(0.985)\lambda^p$. Importantly, a lower demand elasticity argues against a reduced rate on distributional grounds—because it implies a higher revenue loss from cutting the rate, and hence a larger reduction in public spending. These, however, are all (very) special cases; the indirect revenue impact is likely to be highly context-specific and general speculation dangerous.

The significance of $\Gamma_G$ is similar. To the extent that increased public provision changes commodity demands in such a way as to increase tax revenue, so that $\Gamma_G > 0$, this strengthens the case for an increased tax on commodity $k$. Again, however, the sign of this term is ambiguous: increased public spending on health or education, for instance, presumably enhances the capacity to work and so tends, in due course, to increase tax revenue; but it may also weaken the incentive to work in order to pay for privately-provided substitutes for what is now more fully provided by the public sector.

The framework underlying condition (1) is of course very rudimentary. It does not capture, for instance, forms of public spending that make receipt of benefit conditional on some particular actions (such as participation in a work program, or sending children to school); nor does it capture potential complexities arising from the interplay between public and private provision of health and education services. It does, however, point toward the kinds of targeting and incentive concerns that need to be addressed in the joint design of tax and spending systems.

It should be noted too that a range of practical concerns often blunt the impact of subjecting to VAT items that loom large in the budgets of the poorest. Typically, charging VAT is only compulsory for firms beyond some minimum size (in terms of threshold); meaning that smaller firms will escape the tax on their own value added (though of course liable to VAT, without credit or refund, on their own purchases). To the extent that the poor tend to buy from smaller retailers, the prices they face are thus likely to be less affected by the VAT than might at first sight appear.[18] In the case of the Dominican Republic, for which data identifying where purchases are made are available, Jenkins, Jenkins, and Kuo (2006) find that this effect can indeed substantially reduce the regressivity of the VAT. Simple noncompliance with the VAT is likely to have a similar effect.

---

[17] Still assuming $\Gamma_G = 0$.

[18] Even if the tax benefit is retained by small retailers rather than passed on to their customers, the distributional impact is still benign if those retailers are themselves of relatively low income.

**But…**

While the formalities have been cast in the maximin terms of protecting the interests of the very poorest, the harsh truth always arises that, in practice, there will be some that cannot be fully protected against reforms of the kind considered here: not all entitled to some benefits may actually receive it, and it will be hard, for instance, to compensate the childless, healthy, and poor aged, for instance, by increased spending on health care and education. Policy, ultimately, has almost inevitably to look to some wider good.

What ultimately stands out, however, is the singular difficulty not of explaining the case for combined tax-spending reforms of the kind considered above—the essence is often fully understood by policymakers—but of persuading governments to adopt them. This is presumably because they believe they cannot carry their support base with them. That, in turn, may reflect a common sense appreciation that the pain of the increased tax may not always be matched by offsetting gains on the spending side. Or there may be a perception that there are other and even more progressive ways in which to finance that same increase in public spending—by paying for a reduction in the tax on *k* by raising that on some other commodity, for instance. The same issue still arises, however, once maximal progressivity has been achieved in these other financing instruments. Or it may simply be that the richer groups understand precisely how reduced VAT rates act to their particular advantage. It remains unclear whether the failure of policymakers to act on arguments demonstrating that reduced rates are dominated, in equity terms, by other policies is due to weaknesses of the policymaking process or to a political naivety of arguments like those above, which further thought and work might overcome.[19]

[19] There has recently been some resurgence, for example, of the notion that increased use of earmarking might be useful in this respect (Ghana, which has earmarked increased VAT rates to increased health and education expenditure, being a case in point). For references, and a somewhat skeptical assessment, see Keen (2012).

## III. THE ANATOMY OF CASCADING

The Chelliah Report provides an eloquent account of the various dangers of cascading. One is that it makes the structure of the tax system opaque, running counter to basic notions of transparency and simplicity: because the rate at which tax cumulates through the production chain depends on the vagaries of input-output relationships, cascading results in a pattern of effective tax rates that is "…almost fortuitous and largely unknown to policymakers" (p. 36). And because exports carry unrecovered tax on their inputs, it is "…a great hindrance to our export efforts"(p. 38)—and also, by the same token, a great spur to imports, since they enter India free of foreign input taxes, to the extent that VAT there is properly refunded on exports.

The focus here, however, is on the distortion of real production decisions implied by cascading: the effect, that is, of "...unintended changes in the relative prices of inputs and hence in the proportion in which different inputs are used" (p. 37). The consequence of these is that "The increase in consumer prices due to cascading is not limited to what accrues to the Exchequer by way of revenue" (p. 37). That is, the loss to consumer exceeds the revenue raised by the government—a form of deadweight loss. In qualitative terms, the point is widely appreciated. There is surprisingly little work, however, aimed at gauging the potential magnitude of the welfare losses at issue and identifying the precise considerations on which it depends.[20] For that, one must delve deeper into the anatomy of cascading and input taxation.

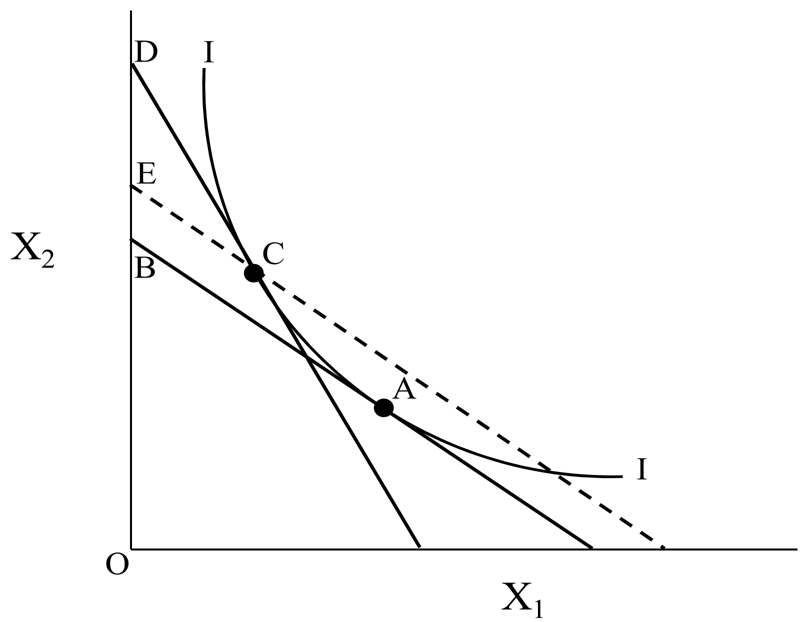
### The deadweight loss from taxing inputs

The fundamental source of loss, in fact, is not cascading. It is, as in the first of the quotations in the preceding paragraph, the distortion of input prices. And this can arise with only one stage of production. Figure 4 illustrates. Here, *II* is the isoquant showing the quantities of inputs $X_1$ and $X_2$ from which one unit of output can be produced. In the absence of taxation, the least-cost input combination at $A$, the point of tangency with an iso-cost line whose slope is (minus) the relative price of input 1; and per-unit production costs, in terms of good 2, are given by the distance $OB$. If now a tax is levied on input 1 (pre-tax prices being assumed constant), making the iso-cost line steeper, the least cost combination becomes that at $C$ (reflecting substitution toward the untaxed input 2) and unit production costs increase to$OD$. Drawing a line through $C$ with slope given by relative pre-tax prices, tax revenue is given by

[20] There appear to have been very few attempts to quantify the distortionary costs from cascading. An early paper by Bovenberg (1987) concluded from a computable general equilibrium model calibrated to the Thailand of the 1970s that raising an additional 1 percent of GDP by means of a uniform rate consumption-type VAT rather than intensifying preexisting cascade taxes would raise real incomes by about 0.004 percent of GDP—a relatively small effect. Focusing on the impact on the cost of investment goods, Smart and Bird (2009) find much stronger effects, estimating that movement toward a VAT in some Canadian provinces raised investment about 12 percent.

$DE$. The point stressed by Chelliah is that this is less than the increase in private unit costs $BD$. The excess of the increase in unit costs over the revenue collected, $ED$, is thus the deadweight loss caused by the input taxation. Equivalently, this deadweight loss is the increase in unit costs induced by the tax when these costs are valued at the pre-tax prices that capture their real social value (as opposed to the private costs perceived by the firm).

**Figure 4. The Deadweight Loss from Input Taxation**




The essence of the Diamond-Mirrlees (1971) theorem is that, under certain conditions, no Pareto efficient tax system can distort input prices in this way: all producers should face the same relative prices, so that relative social and marginal costs of production coincide. This precludes taxing inputs, since that would mean that the price received by the seller is less than that paid by the buyer. Intuitively, distorting production decisions can only reduce aggregate output, which cannot be wise if the elements of that aggregate output can be taxed directly. This theorem rests on a range of assumptions that are unlikely to be met in practice, including an ability to tax all elements of final consumption which, especially in economies marked by widespread noncompliance, is unlikely to be the case. In these circumstances, it may indeed be desirable to tax some transactions between firms[21] and, by the same token, some degree of cascading could be desirable. The concern here, however, is not with the possibility that some degree of input taxation and cascading may be efficiency-improving—

[21] The point is made formally by Newbery (1986). Other conditions for the validity of the Diamond-Mirrlees theorem are nontrivial (requiring, for instance, that any pure profits can be taxed at up to 100 percent and that there be no imperfections of competition), but beside the point for present purposes.

though that is a topic deserving more attention than it has received. Rather it is to understand the nature and potential magnitude of the potential deadweight loss when it is not.

For this, it is useful to start by returning to the example in Figure 4. Expressing the deadweight loss from input taxation shown there as a proportion of the unit production cost in the absence of taxation, $L = EB/OB$, it can be shown that, to a second-order approximation,[22]

$$L \approx \left(\frac{1}{2}\right)\alpha(1-\alpha)\sigma T^2 \tag{2}$$

where $\alpha$ denotes the share of the taxed input in all production costs (in the absence of tax); $\sigma$ is the elasticity of substitution between the two inputs (characterizing the degree of curvature of the isoquants, being larger the flatter they are); and *T* is the tax rate in ad valorem form.[23]

Equation (2) is reminiscent of the familiar Harberger approximation to the deadweight loss from taxing a single item of final consumption, the loss increasing with the square of the tax and (akin to the role of the compensated demand elasticity in the consumption good case) with the ease of substituting away from the taxed and toward the untaxed input. An important limiting case, to which we shall return, is that in which there is no possibility of substitution, so that the isoquant is L-shaped (and $\sigma = 0$): in this case, input taxation does not distort input choices, and deadweight loss is zero.

There is, however, one striking feature of (2). This is that the deadweight loss is not monotonically increasing in the importance of the taxed input in overall production, $\alpha$: indeed the deadweight loss is maximized when its share of total production costs is 50 percent, and decreases as its cost share increases beyond that. The reason is easily seen from Figure 1: if all inputs were taxed at the same proportional rate, then their relative prices to the firm would be unchanged and hence there would be no distortion to the proportions in which they are used. Unit production costs would increase, but by exactly the same amount as revenue to the government increases; so deadweight loss would be zero. The implication,

---

[22] Denote by $C(Q)$ the unit private cost of production as a function of the input price $Q$ faced by the producer, so that the derivative $C'(Q)$ gives the demand for that input per unit of output; and define deadweight loss *DL* as the amount by which the tax-induced increase in private costs exceeds the revenue raised, so that

$$DL = C(P(1+T)) - C(P) - TPC'(Q)$$

where *P* denotes the pre-tax price input price. Taking a second-order approximation in *T*, assuming constant *P*, and dividing by *C(P)* gives (2) of the text, where $\alpha \equiv PC'(P)/C(P)$ and $\sigma \equiv -pCC''/(C - PC')$.

[23] That is, as a proportion of the pre-tax price.

crudely put, is that while taxing inputs is bad, it is not necessarily worse to tax more inputs rather than fewer.

Some sense of the potential magnitude of the deadweight loss implied by (2) is given in Table 2, which reports values of $L$ implied by alternative combinations of tax rate and elasticity of substitution. The input share $\alpha$ is taken throughout to be 50 percent; that is implausibly high (labor costs, for instance, may very well be a larger share of the total input costs than are the commodity inputs at issue here) but the assumption is convenient in giving an upper bound on the deadweight loss[24] at the parameter values shown.

**Table 2. The Deadweight Loss from Input Taxation—Illustrative Calculations**

| Tax Rate (percent) | Elasticity of Substitution | | | | |
|---|---|---|---|---|---|
| | Zero | 0.5 | 1 | 2 | 5 |
| 2.5 | 0 | 0.00 | 0.01 | 0.02 | 0.04 |
| 5 | 0 | 0.02 | 0.03 | 0.06 | 0.16 |
| 15 | 0 | 0.14 | 0.28 | 0.56 | 1.48 |
| 20 | 0 | 0.25 | 0.50 | 1.00 | 2.50 |

Note: Calculated from Equation (2) assuming $\alpha = 0.5$; deadweight loss in percent of undistorted costs.

What emerges is that the deadweight losses become marked only when tax rates and/or the degree of substitutability and/or the tax rate become quite high—reaching 0.5 percent of (undistorted) input costs when $\sigma = 2$, for instance, only when the tax rate is at least 15 percent. In practice, explicit turnover taxes have rarely reached such levels, the received wisdom being that they become unworkable at rates of 10 percent or so. Rates at the higher end of those in Table 2—at which the deadweight loss does seem potentially sizable—are relevant, nonetheless, for two reasons. First, under the VAT itself—often levied at these higher rates—exemptions give rise to unrecovered input taxation, noted in the introduction. Second, the cascading effect of multiple stage production, ignored so far, can imply effective input tax rates far higher than nominal—and it is to this that we turn next.

[24] More precisely, they are upper bounds on the second order approximation.

**Chains of production**

To see the implications of cascading itself, suppose now that there is a single chain of production, comprising $N$ stages (so that in the example of Figure 4 above, $N = 1$).[25] At each stage, each competitive producer spends the same fraction $\alpha$ of its pre-tax costs purchasing, for use as an input, the output of the previous stage (paying the tax-inclusive price), and the remaining $1 - \alpha$ being spent on some untaxed input. For simplicity (and things do become very messy), it is assumed that there is constant return to scale at all production stages, that all goods are sold at tax-inclusive cost, that the input price at the initial stage is fixed, and that all tax-induced cost increases are exactly passed on into producers' selling prices.[26]

***Substitution only at the final stage***

Suppose first that there is no substitution in production at any stage of production except the last. Input taxation at these earlier stages then creates no distortion of input choices at those earlier stages. What it does do, however, is add to tax-inclusive production costs at every stage and so cumulate to amplify the tax-induced increase in input prices faced at the final stage—and so also amplifies the deadweight loss from the distortion of production at the final stage, with the nominal tax rate $T$ in (2) replaced by an effective tax rate $T^e$ that reflects not only the taxation of inputs directly used in that sector but also the taxation of those used to produce these inputs.

[25] In practice, of course, production relationships are much more complex then the literal chain considered here. But the familiar standard input-output framework, in which each good may be used in, and use the products of, many other sectors, does not lend itself so easily to the analysis of substitution between inputs; and the notion of the length of the chain does capture a notion of the complexity or simplicity of production relationships whose relevance to tax design is of interest in itself.

[26] This last assumption is problematic in open economy settings, since full pass-through will not be possible if the same products can be bought on (or can only be sold into) world markets in which producers not subject to similar input taxation are active. In the limiting case of a small open economy (and assuming the tax is levied on imports and remitted at the final stage on exports), tax-induced cost increases would have to be fully passed back to untaxed and internationally immobile inputs—labor being a prime candidate—if domestic production is to continue. This itself simply changes the way in which the inefficiencies discussed here manifest themselves: as reductions in the payments for such inputs (in real labor income, for example) that exceed the tax revenue raised. A full analysis of the open economy case, however, raises further complexities and is not pursued here.

The calculation of the effective rate $T^e$ is cumbersome; details are in Appendix B.[27] Of particular interest is how the effective tax rate varies as the number of prior stages rises. Table 3 reports some illustrative calculations, again assuming $\alpha = 0.5$.

**Table 3. Effective Tax Rates with Cascading—Illustrative Calculations**

| Tax Rate (percent) | Number of Production Stages | | | |
|---|---|---|---|---|
| | 2 | 4 | 6 | ∞ |
| 2.5 | 3.8 | 4.8 | 5.0 | 5.0 |
| 5 | 7.6 | 9.7 | 10.3 | 10.5 |
| 15 | 23.6 | 31.4 | 34.0 | 35.3 |
| 20 | 32.0 | 43.5 | 47.7 | 50.0 |

Note: Figures in percent, and assume $\alpha = 0.5$.

The effective rate of course increases with the number of stages, and since the deadweight loss in (2) increases with the square of the (effective) tax rate, the impact can be sizeable. With three prior production stages, for instance, a nominal rate of 2.5 percent becomes an effective rate of 5 percent—which, from Table 2, triples the deadweight loss. So long as $\alpha(1+T) < 1$, however, the effective rate remains finite as the number of stages goes to infinity, converging to $T/(1-\alpha(1+T))$: intuitively, the output of any stage accounts for only a proportion $\alpha^s$of the input $s$ stages later, so that the effect of an increase in its price vanishes. And an important lesson from Table 3 is that it vanishes quite quickly. A nominal rate of 5 percent, for instance, reaches 98 percent of its asymptotic value with only three prior stages (and, not shown, 85 percent with only two). The point is of some importance, notably for many simpler and less advanced economies seen as having relatively short production chains—that, these results imply, cannot be taken as implying that cascading is of little concern.

[27] The result established there, and underlying Table 3, is that the effective rate of input taxation at the *N*th stage, reflecting input taxation at that and all previous stages, is given by

$$T_N^e(T) = (1+T)(1-\alpha)\left[\frac{1-[\alpha(1+T)]^{N-1}}{1-\alpha(1+T)}\right] + \alpha^{N-1}(1+T)^N - 1.$$

***Substitution at all stages***

This brings us, finally, to the most general and interesting case: that in which input substitution may arise at each of the many stages of production. This introduces additional considerations pointing in opposite directions. Now inefficiency can arise at each stage—presumably amplifying the aggregate welfare loss. At the same time, however, the possibility of substituting away from taxed inputs at one stage blunts the consequent increase in costs, dampening the increase in the price charged to those purchasing that item for use as an input and so reducing the consequent distortion of input choices at that subsequent stage—an effect that tends to reduce the aggregate loss. Clearly a wide range of possibilities arise, depending for instance on how the ease of substitution varies across stages of production.

The analytics now become even messier. Taking, however, the case in which there are just two stages of production and the elasticity of substitution is the same at each, a second-order approximation to the aggregate welfare loss is derived in Appendix C.[28] This again, it is worth noting, has the property that increasing the share of tax inputs will tend, beyond some point, to actually reduce deadweight loss (though a 50 percent share is no longer the turning point).[29] Using this approximation, Table 4 reports results for the case of a nominal tax rate of 15 percent and, for comparison with earlier calculations, a 50 percent share of taxed inputs at both stages.

**Table 4. Deadweight Loss when Input Distortion Arises at Two Stages**

| Elasticity of Substitution | | | | |
|---|---|---|---|---|
| **First Stage** | **Second Stage** | | | |
| | 0.5 | 1 | 2 | 5 |
| 0.5 | 0.39 | 0.70 | 1.34 | 3.23 |
| 1 | 0.46 | 0.77 | 1.41 | 3.30 |
| 2 | 0.60 | 0.91 | 1.55 | 3.45 |
| 5 | 1.02 | 1.34 | 1.97 | 3.87 |

Note: Assumes nominal tax rate of 15 percent, and $\alpha = 0.5$ at both stages.

---

[28] In obvious notation, the approximation derived there is

$$L \approx \left(\frac{1}{2}\right)\alpha_2\{\sigma_2(1-\alpha_2)(1+\alpha_1)^2 + \sigma_1\alpha_1(1-\alpha_1)\}T^2 .$$

[29] When both the elasticity of substitution $\sigma$ and the input share $\alpha$ are the same at both stages, for example, deadweight loss is maximized at $\alpha = 0.61$.

The first of the effects identified above clearly dominates, in these calculations at least, with higher elasticities implying greater losses. And the effects are quite sizeable. Comparing with Table 2 and allowing for an elasticity of substitution of 0.5 at a single prior stage as well as at the last stage, for example, deadweight loss nearly triples.

The numbers presented in Table 4 are of course no more than illustrative (and those in Table 2, it should be recalled, are upper bounds on welfare losses). Clearly, additional work would be needed to calibrate them to plausible circumstances. But, at the very least, it is evidently not hard to construct examples in which—as the Chelliah Report so eloquently argued—the welfare losses associated with cascading are large enough to warrant close attention from policymakers.

**Vertical integration**

All this ignores the possibility that firms may seek to avoid tax on their inputs by vertically integrating: that is, by producing inputs in-house.[30] This has indeed always been a major concern in turnover tax systems, and arises too in relation to VAT exemptions (banks, for instance, having an incentive to provide their own security operations rather than buying them externally and incurring unrecovered tax). This eliminates the distortions addressed above—but introduces another distortion that could potentially be even larger. For such avoidance creates its own deadweight loss, equal—since tax collected falls to zero—to the additional private costs associated with the avoidance activities (developing, for instance, an expertise in security arrangements).[31] Since this avoidance will be privately profitable so long as those additional private costs are less than the tax that would be paid if the inputs were purchased externally, the consequent deadweight loss could be equal to as much as (just less than) the total input tax that would otherwise be collected. Suppose, to take an extreme case, that there is no substitution in input use at any stage. Then firms still have an incentive to vertically integrate to avoid these input taxes, and incur private costs that are a social deadweight loss in doing so, even though the deadweight loss were they to pay the tax would be zero.

[30] Attempts are sometimes made to close avoidance of this kind by charging tax on such self-supplies.

[31] If there were no such costs, these activities would presumably happen without any tax incentive.

## IV. CONCLUDING

The lessons learned above are easily summarized. First, assessing the distributional case for reduced VAT rates requires understanding the distributional impact of the public spending that a higher rate could enable—and even poorly targeted spending may be a better way to support the poor than a reduced rate. Second, output losses from cascading—or more precisely, the amplification of the production inefficiencies to which input taxation can lead—may actually be lower the wider the set of inputs that are taxed; but, probably more to the point, may plausibly be large even at a fairly low nominal tax rate and with relatively few stages of production.

All this is somewhat removed from current GST controversies in India. These, indeed, turn less on the issues addressed above than the difficulties of integrating the treatment of goods and services—what Chelliah had in mind as a "full-fledged" VAT—given the constitutional restriction of powers over each to dint levels of government and the difficulty of bringing trade between states within a federation into the VAT without burdensome fiscal frontiers between them. The aim here has not been to cast new light on these challenges, but rather to bring to bear on two long-standing issues in indirect taxation some of the rigor of thought that marked the work of Raja Chelliah.

**Appendix A. Derivation of Equation (1)**

Suppose there are *H* individuals, the typical member having preferences described by an indirect utility function $V^h(Q_1, \ldots, Q_M, w, Y, g^h)$ defined over the consumer prices $Q_i$ of *M* commodities, wage rate *w* (assumed or normalized to be untaxed and unchanging), lump-sum income *Y* (very possibly zero), and the amount of public spending $g^h = S^h G$ going to their benefit. This spending is financed only by commodity tax revenue, so that $G = \sum_{h=1}^{H} T_i \sum_{i=1}^{M} C_i^h$, where $T_i$ is the tax on *i* (in ad valorem form) and $C_i^h$ is $h$'s consumption of $i$.[32]

Assuming all producer prices to be unchanged, the effect on *p*'s welfare of an arbitrary change in $T_k$ and *G* is given, using Roy's identity, by

$$dV^p = -V_Y^p C_k^p + V_g^p S^p dG\ , \qquad \text{(A.1)}$$

the subscripts to $V^p$ indicating derivatives. Perturbing the government's budget constraint gives

$$dG = \left(\sum_{h=1}^{H} C_k^h + \sum_{h=1}^{H} T_i \sum_{i=1}^{M} (\partial C_i^h / \partial Q_k)\right) dT_k + \left(\sum_{h=1}^{H} T_i \sum_{i=1}^{M} (\partial C_i^h / \partial G)\, S^h\right) dG. \quad \text{(A.2)}$$

Solving (A.2) for *dG* and substituting into (A.1) gives (1) of the text, where $\lambda^p \equiv (V_g^p / V_Y^p)$, $\Gamma_k \equiv \sum_{h=1}^{H} T_i \sum_{i=1}^{M} (\partial C_i^h / \partial Q_k) / \sum_{h=1}^{H} C_k^h$ and $\Gamma_G \equiv \sum_{h=1}^{H} T_i \sum_{i=1}^{M} (\partial C_i^h / \partial G) S^h$.

[32] For simplicity, all consumption is assumed to be taxed, but the analysis is easily generalized.

**Appendix B. Derivation of the Effective Tax Rate $T^e$**

The analysis assumes that, at every stage of production, one unit of output requires $(1-\alpha)$ units of some untaxed input, the price of which is normalized at unity, and $\alpha$ units of an input that is taxed, at all stages, at the ad valorem rate *T*; normalizing the pre-tax price of the untaxed input at the first taxed production stage also to unity, $\alpha$ can also be thought of as the share of the taxed input in unit costs when the tax rate is zero. The price charged by firms at stage *n* is denoted by $P_n$; this differs from the tax-inclusive price paid by firms at the next stage $n+1$ by the amount of the input tax, so that the tax-inclusive input price at stage *n* is given by $Q_{n+1} = P_n(1+T)$. Producers are assumed to make zero profit at each stage; this, together with an assumption that the pre-tax price of the input at the first taxed stage is fixed, means that input taxes charged at each stage are fully passed on in the price charged to firms at the next.

With these assumptions, the price charged by producers at the first stage—the first, that is, at which inputs are taxed—is given by

$$P_1 = 1-\alpha+\alpha Q_1 = 1-\alpha+\alpha(1+T) \qquad \text{(B.1)}$$

and so the tax-inclusive input price faced at the second stage is

$$Q_2 = P_1(1+T) = (1-\alpha)(1+T) + \alpha(1+T)^2\,. \qquad \text{(B.2)}$$

These second stage producers then charge

$$P_2 = 1-\alpha+\alpha Q_2 = 1-\alpha+\alpha(1+T)P_1 = (1-\alpha)\{1+\alpha(1+T)\} + \alpha^2(1+T)^2 \qquad \text{(B.3)}$$

which in turn implies

$$Q_3 = (1-\alpha)\{1+T+\alpha(1+T)^2\} + \alpha^2(1+T)^3\,. \qquad \text{(B.4)}$$

Proceeding in this way, tax-inclusive input costs at stage *N* are

$$Q_N = (1-\alpha)\sum_{i=1}^{N-1}\alpha^{i-1}(1+T)^i + \alpha^{N-1}(1+T)^N. \qquad \text{(B.5)}$$

Define $T_N^e(T) \equiv (Q_N(T) - Q_N(0))/Q_N(0)$ to be the effective rate of input taxation at stage *N*, denoted $T_N^e(T)$, to the proportionate increase in input costs at stage *N*, reflecting taxation both at stage *N* itself and all prior stages. The normalizations above imply $Q_N(0) = 1$; simplifying in (A.5) then gives the result in the text.

## Appendix C. Deadweight Loss with Substitution at Two Stages

With the output of sector 1 used as input in sector 2, defining deadweight loss to be the amount by which the tax-induced increase in the cost of producing good 2 exceeds the tax revenue raised:

$$DL = C_2(P_1(1+T)) - C_2(1) - C_2'(Q_2)P_1T - C_2'(Q_2)C_1'(1+T)T \qquad \text{(C.1)}$$

where $C_i$ denotes the unit cost function in sector *i* and the tax-exclusive price at which sector 1 buys its in input is normalized at unity. Here the third term is the tax collected on taxing 1's inputs, per unit of final output of 2, and the last is that collected on the taxation of inputs into sector 1 ($C_2'(Q_2)$ being the demand for good 1 per unit of output of good 2, and $C_1'(1+T)$ the demand for 1's input per unit of its output of 1).

Noting that $P_1 = C_1$ and collect terms, (C.1) becomes

$$DL(T) = C_2[C_1(1+T).(1+T)] - C_2(1) - C_2[C_1(1+T).(1+T)]\{C_1(1+T) + C_1'(1+T)\}T$$

from which, taking a first order approximation and canceling some terms,

$$DL'(T) = -C_2''\big(C_1.(1+T)\big)\big\{C_1 + C_1'\big\}\big\{C_1 + C_1'T\big\}T - C_2'\big(C_1.(1+T)\big)(C_1'')^2 T \qquad \text{(C.2)}$$

so that evaluating at $T = 0, DL'(0) = 0$. Differentiating again

$$DL''(0) = -C_2''(C_1 + C_1')^2 - C_2''(C_1'')^2 \qquad \text{(C.3)}$$

Multiplying (C.3) by $T^2/C_2$ gives the approximation reported in the text.

**References**

Atkinson, Anthony B., and Joseph E. Stiglitz, 1976, "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, Vol. 6, pp. 55–75.

Barreix, Alberto, Martin Bès, and Jerónimo Roca, 2012, "Solving the Impossible Trinity of Consumption Taxes: Personalized VAT," mimeo, Inter-American Development Bank.

Bird, Richard, and Pierre-Pascal Gendron, 2007, *The VAT in Developing and Transitional Countries* (Cambridge University Press: Cambridge).

Bovenberg, A. Lans, 1987, "Indirect Taxation in Developing Countries: A General Equilibrium Approach." *Staff Papers*, International Monetary Fund, Vol. 34, pp. 333–73.

Castañeda, Tarsicio, 2005, "Targeting Social Spending to the Poor with Proxy-Means Testing: Colombia's SISBEN System," World Bank, Social Protection Discussion Paper Series No. 0529.

Chelliah, Raja J., and others, 1992, *Final Report of the Tax Reforms Committee* (New Delhi: Directorate of Publications, Ministry of Development).

Coady, David, 2004, "Designing and Evaluating Social Safety Nets: Theory, Evidence and Policy Conclusions," International Food Policy Research Institute, FCND Discussion Paper No. 172.

Cnossen, Sijbren, 2012, "Introducing a Modern GST in India: Need for Further Debate," *Tax Notes International*, June 25, 2012.

Crawford, Ian, Michael Keen, and Stephen Smith, 2010, "VAT and Excises," pp. 275–362 in James Mirrlees and others (eds), *Dimensions of Tax Design: The Mirrlees Review*, (Oxford: Oxford University Press for Institute for Fiscal Studies).

Deaton, Angus, 1979, "Optimally Uniform Commodity Taxes," *Economics Letters*, Vol. 2, pp. 357–61.

Diamond, Peter, and James Mirrlees, 1971, "Optimal Taxation and Public Production I: Production Efficiency," *American Economic Review*, Vol. 61, No. 1, pp. 8–27.

Ebrill, Liam, Michael Keen, Jean-Paul Bodin, and Victoria Summers, 2001, *The Modern VAT* (Washington: International Monetary Fund).

Kaplow, Louis, 2006, "On the Undesirability of Commodity Taxation Even When Income Taxation is Not Optimal," *Journal of Public Economics*, Vol. 90, pp. 1235–50.

Keen, Michael, 2012, "Taxation and Development—Again," forthcoming in Clemens Fuest and George Zodrow (eds), *Studies of Critical Issues in Taxation and Development* (Boston: MIT Press).

Jenkins, Glenn P., Hatice P. Jenkins, and Chun-Yan Kuo, 2006, "Is the VAT Naturally Progressive?" (mimeo, Queens University).

Laroque, Guy R., 2005, "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof," *Economics Letters*, Vol. 87, pp.141–44.

Mahal, Ajay, Abdo S. Yazbeck, David H. Peters, and G.N.V. Ramana, 2007, *The Poor and Health Service Use in India* (Washington: World Bank).

Muñoz, Sonia, and Stanley Sang-Wook Cho, 2004, "Social Impact of a Tax Reform: The Case of Ethiopia," pp. 353–84 in Gupta, Sanjeev, Benedict Clements and Gabriela Inchauste (eds), *Helping Countries Develop: The Role of Fiscal Policy* (Washington: International Monetary Fund).

Newbery, David M., 1986, "On the Desirability of Input Taxes," *Economics Letters*, Vol. 20, pp. 267–70.

Organisation for Economic Cooperation and Development, 2007, "Putting Public Finances on a Firmer Footing," Chapter 2 in *OECD Economic Survey: Mexico 2007* (Paris: Organisation for Economic Cooperation and Development).

Poddar, Satya, and Morley English, 1997, "Taxation of Financial Services under a Value-Added Tax: Applying the Cash-flow Method," *National Tax Journal,* Vol. 50, pp. 89–111.

Poddar, Satya, and Ehtisham Ahmad, 2009, "GST Reforms and Intergovernmental Considerations in India," Government of India, Ministry of Finance, Department of Economic Affairs, Working Paper 1/2009.

Rao, M. Govinda, 2011, "Goods and Services Tax: A Gorilla, Chimpanzee or a Genus like Primates"? *Economic & Political Weekly*, Vol. 46, pp. 43–48.

Ring, Raymond J. Jr., 1989, "Proportion of Consumers' and Producers' Goods in the General Sales Tax," *National Tax Journal*, Vol. 42, pp. 167–79.

Sah, Raj K., 1983, "How Much Redistribution is Possible through Commodity Taxes?" *Journal of Public Economics*, Vol. 20, pp. 89–101.

Smart, Michael, and Richard M. Bird, 2009, "The Impact in Investment of Replacing a Retail Sales Tax with a Value-Added Tax: Evidence from Canadian Experience," *National Tax Journal*, 62, pp. 591–609.