The Theory of Optimal Taxation

Sørensen, Peter Birch

Publication date:
2006

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Sørensen, P. B. (2006). *The Theory of Optimal Taxation: What is the Policy Relevance?* Cph.: Economic Policy Research Unit. Department of Economics, University of Copenhagen.

EPRU Working Paper Series

Economic Policy Research Unit
Department of Economics
University of Copenhagen
Studiestræde 6
DK-1455 Copenhagen K
DENMARK
Tel: (+45) 3532 4411
Fax: (+45) 3532 4444
Web: http://www.econ.ku.dk/epru/

The Theory of Optimal Taxation:
What is the Policy Relevance?

Peter Birch Sørensen

2006-07

ISSN 0908-7745
The activities of EPRU are financed by a grant from
The National Research Foundation

First draft, August 2006

THE THEORY OF OPTIMAL TAXATION: WHAT IS THE POLICY RELEVANCE?

Peter Birch Sørensen

Department of Economics, University of Copenhagen, EPRU og CESifo

Invited Plenary Lecture prepared for the 62nd congress of the

International Institute of Public Finance in Paphos, Cyprus, August 28-31, 2006

Abstract: The paper discusses the implications of optimal tax theory for the debates on uniform commodity taxation and neutral capital income taxation. While strong administrative and political economy arguments in favor of uniform and neutral taxation remain, recent advances in optimal tax theory suggest that the information needed to implement the differentiated taxation prescribed by optimal tax theory may be easier to obtain than previously believed. The paper also points to the strong similarity between optimal commodity tax rules and the rules for optimal source-based capital income taxation.

Keywords: optimal taxation, uniform taxation, tax neutrality

JEL codes: H21, H25

Correspondence to:

Peter Birch Sørensen

Department of Economcs, University of Copenhagen

Studiestraede 6, 1455 Copenhagen K, Denmark

E-mail: peter.birch.sorensen@econ.ku.dk

THE THEORY OF OPTIMAL TAXATION:
WHAT IS THE POLICY RELEVANCE?

Peter Birch Sørensen[1]

1. Introduction

The breakthrough of the modern theory of optimal taxation in the early 1970s opened up a new fertile area of research, but it also created a larger communication gap between theorists and practitioners of public finance. To many applied economists working for governments and international organizations, the new theories of optimal taxation seemed highly technical and abstract, and hence of little policy relevance. Even today it is a widespread view that optimal tax theory has produced very few robust results that can serve as a basis for concrete useful policy advice.

This paper argues that the theory of optimal taxation does in fact provide many important lessons for policy makers and that recent theoretical progress in this area may help to bridge the gap between academic research and practical policy advice. At the same time I shall argue that optimal tax theory still has obvious limitations and that many of the practitioners' objections against it should be taken quite seriously.

The theory of optimal taxation is normative, essentially assuming that policy is made by a benevolent dictator who respects individual preferences as well as some 'social' preference for equality. One can choose to dismiss this body of theory by pointing out that actual policy makers typically represent specific interest groups and that actual policies tend to reflect some compromise between conflicting interests rather than the maximization of a Bergson-Samuelson social welfare function. Indeed, this is why models of Public Choice and Political Economy help us to understand what is going on in the real world. But one could likewise dismiss models of competitive markets by pointing out that the Walrasian auctioneer does not exist and that many economic agents have market power. Yet few if any economists would deny that the theory of perfect competition and

[1] Peter Birch Sørensen is Professor of Economics at the University of Copenhagen and chairman of the Danish Economic Council, an advisory body to the Danish government and parliament. He wishes to thank Henrik Kleven Jacobsen and Claus Thustrup Kreiner for stimulating discussions on the topic of this paper. The usual disclaimer applies.

the First Theorem of Welfare Economics provide a useful benchmark for evaluation of resource allocation in actual market economies. In a similar way, assuming that one accepts its philosophical foundations in utilitarianism and methodological individualism, optimal tax theory provides a benchmark against which to evaluate actual public policies. I would also argue that so-called 'naive' advice based on normative economic theory does have some influence on actual policies, although to different degrees in different countries and time periods. After all, many governments and international organizations employ armies of economists brought up on normative welfare economics, and arguments and ideas do have an impact on public policy debates. So even if one's sole ambition is to understand why certain policies are adopted whereas others are not, it would be a mistake to rule out that advice based on normative economic theory could influence the actual course of events.

The literature on optimal taxation is vast, so my discussion will have to be selective.[2] I will focus on the implications of optimal tax theory for a broad issue that has long been the subject of controversy among economists and policy makers. The issue is whether taxes should be uniform and 'neutral' or whether - even in the absence of externalities - they should systematically discriminate between different economic activities? In the latter case, does optimal tax theory offer any useful advice on the proper differentiation of tax rates, not just in qualitative but also in quantitative terms? In particular, do governments have the information and the administrative capacity to implement the tax rules prescribed by optimal tax theory?

The debate on uniformity and neutrality in taxation involves indirect as well as direct taxation. The question whether indirect taxes should be uniform or differentiated has already received a lot of attention in the literature, especially in the early years following the breakthrough of optimal tax theory (see, e.g., Atkinson and Stiglitz (1972, 1976), Sandmo (1974, 1976), and Sadka (1977)). Today the *theoretical* case for differentiated commodity taxes seems widely accepted, but at the same time there is a widespread feeling that governments do not have the information needed to determine the optimal tax rates on specific goods and services so that, on administrative grounds, a case can be made for uniform commodity taxation. However, this paper will argue that once one

[2] For some recent comprehensive surveys, see Auerbach and Hines (2002) and Salanié (2003).

accounts explicitly for the coexistence of household production and market production, it becomes easier to identify the specific commodities that are candidates for special treatment under an optimal indirect tax system.

In the area of direct taxation the predominant view is that taxes on (income from) capital and labor should be uniform or 'neutral'. The issue whether neutrality in direct taxation is actually desirable seems to have attracted relatively little attention in the literature, perhaps because the fundamental Production Efficiency Theorem of Diamond and Mirrlees (1971) established a presumption in favor of neutral taxation. Instead, much of the literature on capital income taxation has tended to focus on how the tax system can be designed to achieve neutrality. This paper argues that because of the growing international mobility of capital in recent decades, the case for neutrality in capital income taxation is no longer so strong as it may have been in earlier times.

I start out in section 2 by offering a selective review of the uniform-tax controversy on indirect taxation. Section 3 then discusses the desirability of 'neutral' direct taxation, focusing on capital taxation. The final section 4 summarizes the main conclusions of the paper.

2. Indirect Taxation: The Uniform-Tax Controversy

It is generally accepted that there is a good case for selective Pigovian taxes (subsidies) on commodities whose production or consumption generate negative (positive) externalities, and most governments do in fact impose excises on the consumption of alcohol, tobacco, gasoline, etc.[3]

There is much less agreement whether, as a matter of practical policy, indirect taxes should be systematically differentiated even in the absence of externalities. The optimal tax revolution in the early 1970s and the introduction of value-added taxation in many countries around the same time led to renewed interest in this issue. Building on Ramsey's classical contribution (Ramsey (1927)), many academics pointed out that a uniform value-

[3] Pigou (1920) derived the optimal level of externality-correcting excises in the absence of other market distortions. Building on the work of Sandmo (1975), recent developments in optimal tax theory has significantly improved our understanding of the factors determining the optimal level of Pigovian taxes in the presence of other distortionary taxes. See, e.g., Bovenberg and Goulder (1996) and Sandmo (2000).

added tax was very unlikely to be optimal. In the opposite camp many practitioners of public finance argued for uniform taxation. Although the principles of Ramsey taxation are familiar, it is useful to briefly restate them as a basis for the subsequent discussion of the arguments for and against differentiated indirect taxation.

2.1. Ramsey Taxation

It is well-known that a uniform ad valorem tax on all forms of consumption - including the consumption of leisure - would work like a non-distortionary lump sum tax on the value of the consumer's exogenous time endowment. But in practice governments cannot observe and tax the consumption of leisure, so any real-world tax system will tend to cause distortionary substitution towards leisure. It is equally well-known that a uniform ad valorem tax on all commodities other than leisure would be equivalent to a proportional tax on labor income. Whether indirect taxes should be differentiated is thus equivalent to asking whether the labor income tax should be supplemented by selective commodity taxes.

Consider a simple setting with a representative household consuming goods (G), services (S), and leisure (L), enjoying utility

$$U = U(G, S, L), \qquad L = E - N, \tag{2.1}$$

where N is the time spent working in the labor market, and E is the total time endowment. The consumer's budget constraint is

$$P_G G + P_S S = WN, \qquad P_G = p_G + t_G, \qquad P_S = p_S + t_S, \tag{2.2}$$

where P_G and P_S are consumer prices of the two commodities, p_G and p_S are (fixed) producer prices, t_G and t_S are excise tax rates, and W is the consumer price of leisure, that is, the after-tax wage rate (adjusted for any uniform indirect ad valorem tax). For the moment, let us ignore the labor income tax and choose leisure as our numeraire good, setting $W = 1$. The consumer's indirect utility function may then be written as $V = V(P_G, P_S)$, and total government revenue (R) becomes

$$R = t_G G + t_S S. \tag{2.3}$$

The optimal commodity tax problem is to maximize consumer utility for any given amount of revenue collected or, equivalently, to maximize revenue for any given utility level. Using Roy's identity and the symmetry properties of the Slutsky matrix, the solution to this problem implies

$$\frac{t_G}{P_G}\varepsilon_{GG} + \frac{t_S}{P_S}\varepsilon_{GS} = \frac{t_G}{P_G}\varepsilon_{SG} + \frac{t_S}{P_S}\varepsilon_{SS}, \tag{2.4}$$

where the ε-variables are *compensated* own-price and cross-price elasticities of demand for the two commodities. Equation (2.4) states the familiar Ramsey principle that (at the margin) the optimal commodity tax system causes an equi-proportionate reduction of the compensated demands for all goods and services. In other words, the optimal tax system distorts *quantities* as little as possible; it does not necessarily avoid changes in relative commodity prices.

Since the compensated demand functions are homogeneous of degree zero, one can rewrite (2.4) as[4]

$$\frac{t_S/P_S}{t_G/P_G} = \frac{\varepsilon_{GG} + \varepsilon_{SS} + \varepsilon_{SL}}{\varepsilon_{GG} + \varepsilon_{SS} + \varepsilon_{GL}}, \tag{2.5}$$

where ε_{GL} and ε_{SL} are the compensated cross-price elasticities between leisure and the demand for the two commodities. Equation (2.5) is the famous Corlett-Hague rule stating that the commodity which is more complementary to (less substitutable for) leisure should carry a relatively high tax burden in order to offset the tendency of the tax system to induce substitution towards leisure (Corlett and Hague, 1953). In our simple setting uniform taxation is optimal only in the special case where goods and services are equally substitutable for (complementary to) leisure.

This analysis includes only three goods and abstracts from consumer heterogeneity. Christiansen (1984) considered which commodity taxes should supplement the income tax in an economy with many commodities and a continuum of heterogeneous consumers with different exogenous levels of labor productivity. Assuming that the government is concerned about equity as well as efficiency and that it optimizes the non-linear labor income tax, Christiansen found that a commodity should be taxed (subsidized) if it is positively (negatively) related to leisure in the sense that more (less) of the good is consumed if more leisure is obtained at constant income. This result clearly has the same flavor as (although it is more general than) the Corlett-Hague rule: the indirect tax system

[4] Specifically, I use the facts that $\varepsilon_{GG} + \varepsilon_{GS} + \varepsilon_{GL} = 0$ and $\varepsilon_{SS} + \varepsilon_{SG} + \varepsilon_{SL} = 0$ to eliminate ε_{GS} and ε_{SG} from (2.4).

should discourage the purchase of commodities that tend to be consumed jointly with leisure. Extending the analysis of Christiansen (op.cit.) to a setting with heterogeneous consumer tastes, Saez (2002) showed that the optimal non-linear labor income tax should be supplemented not only by excises on commodities that are consumed jointly with leisure, but also by excises on commodities for which high-income earners tend to have a relatively strong taste.

2.2. The Case For Uniform Commodity Taxation

The classical analyses by Ramsey and Corlett and Hague and their modern generalizations would thus seem to provide a strong case for non-uniform commodity taxation. But these studies also point to an obvious practical obstacle to the implementation of an optimal commodity tax system: very little is known about the size and even about the sign of the compensated cross-price elasticities between leisure and all the various goods and services, so the empirical basis for differentiating indirect taxes is very weak. Based on the principle of insufficient reason, one could therefore argue that tax policy makers should act as if all commodities were equally substitutable for leisure. Returning to our three-good framework, this would imply that the utility function (2.1) takes the weakly separable form

$$U = U\left(u\left(G, S\right), L\right) \tag{2.6}$$

where $u\left(G, S\right)$ is a subutility function. Assuming such separability in preferences (but allowing for many goods), Atkinson and Stiglitz (1976) showed that it is inoptimal to differentiate taxes across commodities if the government can use a non-linear labor income tax to achieve its distributional goals. The intuition for this important result is clear: when all commodities are equally substitutable for leisure, there is no second-best efficiency case for distorting the choice between them in order to offset the labor-leisure distortion. Nor is there any equity case for imposing excises, since a labor income tax is a better-targeted instrument for redistribution in a world where innate differences in labor productivity are (assumed to be) the only source of inequality.

As a supplement to this theoretical argument in favor of uniform taxation, practitioners and policy advisers typically stress three other points. The first one is that a uniform VAT is much easier to administer and much less susceptible to fraud than a VAT system with several differentiated rates. In practice this is undoubledly a strong argument in favor of uniformity. The second point is that a commodity tax system differentiated according to Ramsey principles would require frequent changes in tax rates in response to changes in tastes and technologies. This would introduce an extra element of risk and uncertainty into the tax system which might hamper long-term planning and investment. A third point is that acceptance of differentiated taxation as a general principle might invite special interest groups to lobby for low tax rates on their particular economic activities. From a political economy perspective, adherence to a principle of uniformity may therefore provide a stronger bulwark against wasteful lobbyism.

Considering the lack of solid evidence on compensated cross price elasticities with leisure as well as the administrative and political economy arguments against differentiated taxation, there appears to be a strong case for uniformity in indirect taxation, except for areas with an obvious need for internalization of externalities. However, as recent contributions to optimal tax theory have shown, once one allows for household production, the case for uniform taxation is weakened considerably.

2.3. Optimal Commodity Taxation With Household Production

While productive activities within the household sector may involve the production of goods, they typically take the form of production of services. Indeed, much of the output from household production is a very close substitute for services that may also be delivered from the market. For example, think of housing repair, repair of other consumer durables, child care, cleaning and window-cleaning, garden care, cooking, etc. Let us therefore augment our simple three-good set-up by assuming that the total consumption of services consists of services supplied from the market (S^m) and services produced within the household (S^h) so that

$$S = S^m + S^h, \qquad S^h = h(H), \qquad h' > 0, \quad h'' < 0, \tag{2.7}$$

where H is time spent on household production, and $h(H)$ is a concave household production function. The consumer's utility is still given by (2.1), but the amount of leisure is now equal to

$$L = E - N - H. \tag{2.8}$$

To focus on the differential taxation of services, let us now choose the G-good as our numeraire ($P_G = 1$) to obtain the consumer budget constraint

$$G + P_S S^m = WN, \qquad P_S = p_S + t_S, \qquad W = w(1-\tau), \tag{2.9}$$

where w is the pre-tax wage rate and τ is a labor income tax rate (which might also reflect a possible uniform ad valorem tax on all goods and market-produced services). Note that with this normalization, the excise tax rate t_S reflects the *differential* tax on services relative to the tax on goods.

Using this set-up inspired by Sandmo (1990), and allowing for optimization of the labor income tax as well as commodity taxes, Kleven, Richter and Sørensen (2000) (henceforth KRS) showed that the optimal tax system will reduce the compensated demands for all *market-produced* commodities in equal proportions. This accords with the basic Ramsey principle, but it does imply an important modification of the Corlett-Hague rule. Specifically, KRS show that the optimal tax system must satisfy

$$\frac{t_S}{P_S} = \left(\frac{\tau}{1-\tau}\right)\left[\frac{\varepsilon_{SL} + \left(\frac{S^m}{S}\right)\left(\frac{L}{N}\right)\varepsilon_{LL} - \left(\frac{H}{N}\right)\left(\frac{G}{P_S S}\right)\varepsilon_{HL}}{\varepsilon_{SS} + \left(\frac{P_S S^m}{WN}\right)\varepsilon_{SL} - \left(\frac{H}{N}\right)\left(\frac{G}{P_S S}\right)\varepsilon_{HS}}\right], \tag{2.10}$$

$$\varepsilon_{HL} \equiv \frac{\partial H}{\partial W}\frac{W}{H} < 0, \qquad \varepsilon_{HS} \equiv \frac{\partial H}{\partial P_S}\frac{P_S}{H} > 0, \qquad \varepsilon_{LL} \equiv \frac{\partial L}{\partial W}\frac{W}{L} < 0, \qquad \varepsilon_{SS} \equiv \frac{\partial S}{\partial P_S}\frac{P_S}{S},$$

where ε_{HL} and ε_{HS} are the elasticities of home production with respect to the after-tax market wage and the consumer price of services, respectively. Now suppose that service and leisure are complements ($\varepsilon_{SL} < 0$) and that there is no home production ($H = 0$). Since the compensated own-price elasticities ε_{LL} and ε_{SS} are negative, and assuming realitistically that the income tax rate is positive, we then see from (2.10) that the optimal value of t_S is positive, that is, services should be taxed more heavily than goods. This is just a restatement of the Corlett-Hague rule. But suppose now that home production is positive and sizeable so that S^m/S is considerably below unity and H/N is well above zero. According to (2.10) it is then quite possible that services should be *subsidized* ($t_S < 0$) even if they are complementary to leisure.

The point is that a high tax on complements to leisure may not be an efficient way of stimulating tax-discouraged labor supply to the market when such a commodity tax encourages substitution of home production for market production. Taxes should distort the pattern of *market* activity as little as possible, and since untaxed home production tends to reduce market production of services relative to the market production of other goods - because household production mainly takes the form of services - there is a presumption in favor of a lenient tax treatment of services. Indeed, KRS show that when goods and services are equally substitutable for leisure (entering into a homothetic subutility function) so that uniform taxation would be optimal in the absence of home production, services should definitely be taxed at a lower rate than goods when they can be produced in the household sector as well as in the market.

This analysis is relevant for the current tax policy debate in Europe where several countries have recently experimented with reduced rates of VAT on (or other tax concessions to) a number of labor-intensive services that are easily substitutable for home-produced services. As indicated, optimal tax theory suggests that there may be a rationale for such a policy. The practical applicability of this theoretical result is strenghtened by the fact that it is fairly easy to identify a number of services that are close substitutes for home production (cf. the examples given earlier). Yet, from a policy viewpoint a weakness of the theory is that to implement the optimal degree of tax differentiation, we still need to know a number of elasticities which are hard if not impossible to measure.

However, a recent innovative contribution to optimal tax theory by Kleven (2004) suggests that the information needed for an optimal differentiation of commodity taxes may be easier to obtain than previously thought. Kleven analyzes optimal commodity taxation in the generalized household production framework proposed by Becker (1965) where all utility-generating consumption activities require the combination of some good or service with household time spent on the act of consumption (or on acquiring the good). In the Becker approach, our previous utility function (2.1) would be replaced by

$$U = U\left(Z_G, Z_S\right), \tag{2.11}$$

where Z_G and Z_S are the 'activities' of consuming goods and services, respectively. The utility-generating consumption activities (which may be described as 'household production') require inputs of time as well as commodities, so

$$G = a_G Z_G, \qquad N_G = n_G Z_G, \qquad S = a_S Z_S, \qquad N_S = n_S Z_S, \tag{2.12}$$

where the a's and n's are fixed input coefficients, and N_G and N_S are the amounts of time spent on consuming goods and services, respectively. The consumer also spends an amount of time N working in the market, so her time constraint is

$$N + N_G + N_S = E. \tag{2.13}$$

In addition, the consumer faces the usual budget constraint (2.2).

Within such a setting, allowing for an arbitrary number of different consumption activities and assuming a fixed government revenue requirement R, Kleven (2004) demonstrates that if all consumption activities require some positive commodity input, the optimal *ad valorem* tax rate t_j on commodity j is[5]

$$t_j = \frac{R}{\alpha_j} = \frac{R}{1-\beta_j}, \qquad \alpha_j \equiv \frac{P_j a_j}{P_j a_j + W n_j}, \qquad \beta_j \equiv \frac{W n_j}{P_j a_j + W n_j}. \tag{2.14}$$

Equation (2.14) is a strikingly simple *inverse factor share rule* stating that the optimal tax rate on a given commodity is inversely related to the share of commodity input relative to total factor input required in the relevant household production activity (with the value of inputs being measured at consumer prices). Equivalently, the second equality in (2.14) shows that the larger the time input relative to total factor input into some household activity, the higher is the optimal tax rate on the commodity input into this activity. Thus the optimal tax system imposes relatively high tax rates on commodities whose consumption require a large input of household time. In this way the optimal tax system minimizes the amount of time that is diverted from market work to consumption activity within the household sector.

At a basic level this Becker-inspired approach to optimal taxation conforms with the conventional Ramsey approach: tax policy should strive to minimize tax-induced substitution towards non-taxable uses of time. But the approach suggested by Kleven (op.cit.) also offers new interesting insights. From a theoretical perspective, a fundamental point is that a tax system satisfying (2.14) ensures a *first-best* allocation. To see this, note that by combining (2.2), (2.12) and (2.13) and choosing labor as our numeraire ($W = 1$), the consumer's budget constraint may be written as

$$\sum_j Q_j Z_j = E, \qquad Q_j \equiv P_j a_j + n_j, \qquad j = G, S, \tag{2.15}$$

[5]Kleven follows the usual procedure of normalizing the consumer wage to one.

where Q_j is the consumer price (opportunity cost) of consumption activity j. Since $P_j \equiv p_j + t_j P_j$, and since the optimal tax rule (2.14) implies $t_j P_j a_j = RQ_j$, we have $Q_j = (p_j + t_j P_j) a_j + n_j = (p_j a_j + n_j) / (1 - R)$ so that (2.15) becomes

$$\sum_j q_j Z_j = E(1 - R), \qquad q_j \equiv p_j a_j + n_j, \tag{2.16}$$

where q_j is the fixed producer price of activity j. Thus, although the use of time in the household sector cannot be taxed directly, a commodity tax system satisfying (2.14) is seen from (2.16) to be equivalent to a non-distortionary tax on the consumer's exogenous total time endowment E. To put it another way, a commodity tax system satisfying the inverse factor share rule (2.14) is equivalent to a uniform tax on all market goods *and* household time. To achieve uniformity of taxation of *all* household activities, thereby preserving the first-best, it is thus necessary to differentiate the taxation of commodities in inverse relation to the amount of time required for their consumption.

From a practical policy perspective, an interesting insight from Kleven's analysis is that under the assumptions made above, the optimal tax policy depends solely on observable factor shares rather than on unobservable compensated price elasticities. A combined survey of consumption expenditures and household time allocation would in principle provide the information needed to implement the optimal policy, by enabling policy makers to estimate the factor shares α_j determining the optimal tax rates in (2.14). According to the inverse factor share rule, any type of consumption which uses little time, or even saves time, should carry a relatively low tax rate. Many services have low or even negative time intensities from the consumer's viewpoint: hiring somebody in the market to supply a service rather than engaging in do-it-yourself activities saves household time. For these market-produced services the value of the parameter n_j in (2.14) is very low, so such services should be favored by the tax system, just as implied by the more traditional KRS-model of household production discussed earlier.

Note that the assumption of fixed input coefficients a_j and n_j does not rule out the possibility of substitution in household production, since the utility function allows substitution between different activities requiring different inputs of household time. Kleven (2004, p. 548) gives the example of dishwashing which may be carried out either using a brush or a dishwashing machine. These may be seen as two different activities entering the utility function and requiring different fixed input combinations of time and commodities. Because of such substitution possibilities, the assumption of a Leontieff technology in individual household activities is less restrictive than it may seem.

However, the simple Becker framework above also assumes that all utility-generating activities require a positive input of goods or services. If some activity Z_j constitutes 'pure' leisure, requiring no commodity input at all (i.e., if the coefficient a_j is zero), it is no longer possible to mimic a non-distortionary tax on the consumer's time endowment through a commodity tax system that follows the simple inverse factor share rule (2.14). In the case with pure leisure the optimal tax policy can therefore only achieve a second-best allocation. Kleven (op.cit.) shows that in this case the optimal tax rates will generally depend on the compensated own price and cross price elasticities as well as on the factor shares for the different consumption activities. Thus the problem of obtaining reliable estimates of the unobservable compensated elasticities reemerges.

Still, it is hard to think of quantitatively important uses of household time that do not require some form of commodity input, so the possible existence of pure leisure does not seem to be a serious objection to the Becker-inspired model of optimal commodity taxation. A more relevant concern is that for administrative or other reasons a number of goods and services simply cannot be brought into the tax net. Taxes on the remaining commodities will then inevitably be distortive, so a first-best allocation via a simple inverse factor share rule will be unattainable. Even so, Kleven's analysis suggests that data on the allocation of household time can help policy makers to determine a rational structure of indirect taxation.

In summary, recent research on optimal commodity taxation has provided a stronger basis for policy advice on the design of indirect taxation. In the final section of the paper I will elaborate this point and try to draw a policy conclusion, but before doing so, I will discuss the issue of uniformity versus selectivity in direct taxation, since this involves many of the same problems as those arising in the field of indirect taxation.

3. Direct Taxation: The Debate on 'Neutral' Capital Income Taxation

While public finance economists still debate the proper design of the indirect tax system, there seems to be a lot more agreement that direct taxes should be uniform across different production sectors. In particular, most academics as well as practitioners appear to agree that *if* policy makers wish to tax income from capital, they should do so in a 'neutral' manner, imposing the same effective tax rate on all forms of capital income to avoid distorting the pattern of investment. Despite this typical advice from tax experts, politicians throughout the world have been very reluctant to follow the principle of tax neutrality. Indeed, existing systems of capital income taxation tend to be a jungle of special provisions and exemptions for some forms of capital income coupled with sometimes punitive effective tax rates on other types of income from capital. There are many (bad) political economy reasons for this state of affairs. In this main section I will discuss whether optimal tax theory can also help to explain and justify some of the differentiation of capital tax rates observed in the real world.

A premise for my discussion is that the government has decided to include capital income in the tax base. Several contributions to the optimal tax literature (e.g., Chamley (1986) and Judd (1985)) have suggested that the optimal tax rate on the normal return to capital is in fact zero, but more recent contributions have identified a number of reasons why a benevolent government might want to tax the normal return at a positive rate after all (see, e.g., the surveys by Auerbach (2006) and Sørensen (2006)). The issue remains controversial, but here I simply assume that governments must raise some revenue from capital income taxes. The question then is whether optimal tax theory prescribes a uniform rate of tax on all forms of capital income? I will start by discussing this issue in the context of a closed economy before moving on to the open economy.

3.1. The Case for Tax Neutrality

In contrast to differentiated commodity taxes on final consumption goods, differential capital income taxes are a form of input taxes that generate a *production* distortion, causing the marginal rate of substitution between capital and other production factors to differ across production sectors. In their seminal contribution to optimal tax theory, Diamond and Mirrlees (1971) showed that the optimal second-best tax system avoids such production distortions, provided the government can tax away pure profits and can tax households on all transactions with firms. The intuition for this Production Efficiency Theorem is that when the government confiscates all rents and is able to tax all the market transactions of households, it already controls all of the incomes and prices affecting consumer welfare. Hence it has no second-best motive to add further distortions through input taxes in order to offset pre-existing distortions that it cannot otherwise affect. By contrast, as Stiglitz and Dasgupta (1971) were quick to point out, if pure profits cannot be taxed away, and/or if some household transactions cannot be taxed, it will generally be second-best optimal to use distortive input taxes as an indirect means of taxing pure profits and of taxing consumer goods that cannot be taxed explicitly.

The powerful Diamond-Mirrlees Production Efficiency Theorem undoubtedly helps to explain why so many economists consider neutral capital income taxation to be desirable. To be sure, the assumptions underlying the theorem are restrictive, but the work of Auerbach (1989) suggests that even when they are violated so that tax neutrality is inoptimal, the welfare cost of sticking to neutrality is likely to be small. Based on a calibrated dynamic model of the private U.S. economy, Auerbach estimated that the welfare gain of moving from neutral capital income taxation to a system with optimal differential capital tax rates would be negligible for plausible parameter values.

The principle of tax neutrality can also be defended by a number of other arguments that are rather similar to the practical arguments against differentiated commodity taxation. First, even if differential capital income taxation may be theoretically optimal, we do not have firm empirical evidence on all the substitution elasticities in production and consumption that would be necessary to implement the optimal degree of tax differentiation. Second, the optimal degree of tax differentiation will change with changes in tastes and technology, creating an unstable tax system. Third, differentiating capital income taxation across sectors would require drawing a borderline between the different sectors, inducing firms to reclassify themselves as belonging to tax-favored sectors. Fourth, with differential capital income tax rates across sectors, conglomerate firms operating in several sectors would have ample opportunities to reduce total taxable profit through transfer-pricing. Fifth, accepting differential capital income taxation as a general principle invites special interest groups to lobby for tax concessions.

Taken together, all of these theoretical and practical arguments would seem to constitute a formidable case for neutrality in capital income taxation. But as the next section will argue, this case becomes less clear-cut once we explicitly account for the openness of the economy.

3.2. Is Neutral Capital Income Taxation Desirable in an Open Economy?

To illustrate this, it is useful to set up a simple two-sector general equilibrium model to study the effects of sector-specific capital taxation. The model is in the spirit of Harberger (1962), but unlike him I assume that the economy is small and open to trade in goods and capital. Thus the economy faces given world market prices of goods, and since capital is perfectly mobile internationally, the required rate of return on capital is likewise determined in the world market. Labor is not mobile across borders, but perfectly mobile between the two domestic sectors. To highlight the importance of pure rents for the optimal tax policy, I assume that there is also a third factor such as land which is fixed and immobile between sectors. Normalizing the fixed world price ratio between the two domestically produced goods to unity, and leaving the fixed factor behind the curtain, the income Y_j generated in sector j is then given by the production functions

$$Y_1 = f\left(K_1, N_1\right), \qquad Y_2 = F\left(K_2, N_2\right), \tag{3.1}$$

where K and N are the inputs of capital and labor, respectively, and where the marginal products of the two factors are positive but diminishing. Because of the fixed factor, the production functions are assumed to be strictly concave, displaying decreasing returns to scale in capital and labor.

The government levies a unit tax τ on capital invested in sector 1, a unit tax t on capital invested in sector 2, and a uniform labor tax T on labor employed in both sectors. Capital mobility ensures that investors earn the same after-tax return r whether they invest at home or abroad, while domestic labor mobility enforces a common after-tax wage rate w in the two domestic sectors. Profit-maximizing firms employ factors up to the point where the value of their marginal products are equal to their tax-inclusive prices, implying

$$f_K\left(K_1, N_1\right) = r + \tau, \qquad F_K\left(K_2, N_2\right) = r + t, \tag{3.2}$$

$$f_N\left(K_1, N_1\right) = w + T, \qquad F_N\left(K_2, N_2\right) = w + T, \tag{3.3}$$

where the subscripts K and N indicate derivatives of the production functions with respect to the relevant factors. International capital mobility enforces the arbitrage condition

$$r = r^* - t^*, \tag{3.4}$$

where r^* is the exogenous rate of return on foreign investment, net of any source-based taxes levied abroad, and t^* is a residence-based tax on foreign investment levied by the domestic government. The total supply of labor is fixed at $\overline{N}$, and domestic residents are endowed with a fixed total stock of capital $\overline{K}$, so

$$N_1 + N_2 = \overline{N}, \qquad K_1 + K_2 + K^* = \overline{K}, \tag{3.5}$$

where K^* is the (positive or negative) amount of capital invested abroad.

With fixed factor endowments, undistorted product markets, and identical households, maximizing the welfare of the representative domestic consumer is equivalent to maximization of total national income (I) which is

$$I = Y_1 + Y_2 + r^* K^*. \tag{3.6}$$

The government must raise sufficient revenue to finance the exogenous level of public expenditure R, so the maximization of national income takes place subject to the government budget constraint[6]

$$\tau K_1 + t K_2 + t^* K^* + T\overline{N} = R. \tag{3.7}$$

We also allow for the possibility that, for reasons not explained by the model, the tax burden imposed on labor cannot exceed some exogenous limit $\overline{T}$:

$$T \leq \overline{T}. \tag{3.8}$$

Suppose that this constraint is binding and that $\overline{TN} < R$ so that some amount of revenue has to be raised from taxes on capital. Suppose further that the government is

[6] Without affecting the conclusions drawn below, we could have included a non-distortionary tax on pure rents. Such a tax would be a perfect substitute for the uniform labor tax T which is non-distortionary in the present model where total labor supply is fixed. The only crucial assumption for the results presented below is that the government cannot cover all of its revenue needs by means of non-distortionary revenue sources.

in fact able to tax foreign as well as domestic investment. It is then easy to show that maximization of national income (3.6) subject to (3.7) requires

$$\tau = t = t^* \tag{3.9}$$

$$\Longrightarrow \qquad f_K = F_K = r^*. \tag{3.10}$$

In other words, when foreign investment can be taxed, it is optimal to levy a uniform capital tax on all forms of investment. In this way production efficiency is maintained, and the marginal social returns to all domestic investments (f_K and F_K) are kept equal to the marginal social return to foreign investment (r^*).[7] Note that even though pure profits are not (fully) taxed, production efficiency is still desirable in our model because a uniform residence-based capital tax is equivalent to a non-distortionary lump sum tax on the fixed total capital endowment, i.e., the policy (3.9) preserves the first-best allocation.[8]

Thus openness of the economy does not destroy the case for tax neutrality, provided residence-based capital taxation is feasible. But effective enforcement of the residence principle requires that governments are willing to engage in systematic international information exchange on a multilateral basis, and so far they have been very reluctant to do so.[9] Hence it is difficult and often impossible for the domestic tax authorities to

[7] When the tax is levied on the *income* from capital rather than on the stock of capital itself, maximization of national income is achieved by taxing all capital income at the uniform rate t^i, since capital mobility will then enforce the arbitrage condition

$$f_K\left(1-t^i\right) = F_K\left(1-t^i\right) = r^*\left(1-t^i\right)$$

$$\Longleftrightarrow \qquad f_K = F_K = r^*.$$

Since r^* is measured net of any taxes levied in the foreign source country, this tax policy implies that the domestic government allows a *deduction* for foreign taxes from the foreign source income subject to domestic tax. Thus a deduction system of international double tax relief is optimal from a national viewpoint, as pointed out many years ago by Musgrave (1969).

[8] When savings are endogenous, a residence-based capital income tax no longer preserves the first-best allocation. If pure profits cannot be (fully) taxed, it then becomes optimal to levy a source-based capital income tax as an indirect means of taxing rents, especially if these rents accrue to foreigners (see Huizinga and Nielsen (1997)).

[9] The so-called Savings Directive of the European Union tries to take a first step in the direction of systematic information exchange, but it only covers interest income and a small subset of the many tax jurisdictions in the world. Keen and Ligthart (2006) analyze how the incentives for international information exchange might be improved through revenue sharing between source and residence countries.

monitor capital invested abroad. In many economic analyses it is therefore assumed that capital can only be taxed on a source basis, i.e., the domestic government can only tax capital invested within the domestic economy.

Suppose therefore that the policy instrument t^* is not available, but suppose also for a moment that $\overline{TN} \geq R$ so that no revenue has to be raised from taxes on capital. Then it is easily seen that $\tau = t = 0$ is the optimal policy. In other words, when taxes on the internationally immobile factors of production (labor and land) are sufficient to cover the need for public revenue, no source-based capital taxes should be levied, since one thereby avoids distorting capital away from domestic uses and ensures that the condition for production efficiency (3.10) is met. Even if the supply of the internationally immobile factors were elastic, it would still be optimal to avoid source-based capital taxation, as shown by Razin and Sadka (1991). With perfect capital mobility, a source tax on capital would be fully shifted onto the immobile factors, thus distorting their supply while at the same time distorting the use of capital in the domestic economy. Since the latter distortion does not help to reduce any other distortions but only adds an additional efficiency loss, it should be avoided. This may be seen as an application of the Diamond-Mirrlees Production Efficiency Theorem to the open economy context.

However, suppose realistically that for some (political or other) reason a certain amount of revenue has to be raised from source-based capital taxes, i.e., $\overline{TN} < R$. As shown in the appendix, the optimal tax policy then implies

$$\varepsilon_{\tau}^{K_1} + \varepsilon_{t}^{K_1} = \varepsilon_{\tau}^{K_2} + \varepsilon_{t}^{K_2}, \tag{3.11}$$

$$\varepsilon_{\tau}^{K_1} \equiv \frac{\tau}{K_1}\frac{\partial K_1}{\partial \tau}, \qquad \varepsilon_{t}^{K_1} \equiv \frac{t}{K_1}\frac{\partial K_1}{\partial t}, \qquad \varepsilon_{\tau}^{K_2} \equiv \frac{\tau}{K_2}\frac{\partial K_2}{\partial \tau}, \qquad \varepsilon_{t}^{K_2} \equiv \frac{t}{K_2}\frac{\partial K_2}{\partial t},$$

where the epsilons are the elasticities of capital demand with respect to the sector-specific capital tax rates. Equation (3.11) is a *Ramsey rule for capital taxation* stating that, at the margin, the optimal tax system causes the same relative reduction of investment in the different production sectors. In general, this policy rule calls for *differential* capital taxation, just as the standard Ramsey rule for indirect taxation generally requires non-uniform taxation.

Would policy makers be able to implement the Ramsey rule for capital taxation on the basis of observable variables? To investigate this, let us assume that the production functions in (3.1) are of the Cobb-Douglas form so that

$$Y_1 = K_1^{\alpha} N_1^{\beta}, \qquad \pi \equiv 1 - \alpha - \beta > 0, \tag{3.12}$$

$$Y_2 = K_2^{\widehat{\alpha}} N_2^{\widehat{\beta}}, \qquad \widehat{\pi} \equiv 1 - \widehat{\alpha} - \widehat{\beta} > 0, \tag{3.13}$$

where π and $\widehat{\pi}$ are the pure profit shares (land rents) accruing to the fixed factor in the two sectors. In the appendix I use (3.2), (3.3), and (3.11) through (3.13) to derive the following formula for the optimal relative tax rates on the two sectors:

$$\frac{\tau / f_K}{t / F_K} = \frac{\pi + \beta \left(1 - \widehat{\pi} - \widehat{\beta}\right) + (\pi + \beta)\left(1 - \widehat{\beta}\right)(N_1/N_2)}{\left[\widehat{\pi} + \widehat{\beta}\,(1 - \pi - \beta)\right](N_1/N_2) + \left(\widehat{\pi} + \widehat{\beta}\right)(1 - \beta)}. \tag{3.14}$$

The magnitudes τ/f_K and t/F_K are the marginal effective tax rates on capital *income* generated in the two sectors, and β and $\widehat{\beta}$ are the labor income shares. The optimal tax formula (3.14) thus implies that the optimal relative capital income tax rate on a given sector is higher the higher the pure profit share and the higher the labor income share.

To explain these results, let us focus on sector 1 and note that, ceteris paribus, a higher value of π or β must imply a lower value of α, that is, a lower capital intensity of production. This in turn tends to make investment in both sectors less sensitive to the capital tax rate on sector 1, allowing the government to raise that tax rate without distorting the allocation of capital too much.[10] An alternative way of explaining the role of the pure profit share is to note that a sector-specific capital tax works in part as an indirect tax on the pure profits generated in the sector, because it reduces investment in the sector, thereby curbing the demand for the fixed factor used in the sector. In itself a tax on pure profits is non-distortionary, so the greater the relative importance of rents in a sector, the less distortionary is a capital tax on that sector.[11]

[10] Specifically, if we denote the cost of capital in sector 1 by $\rho_1 \equiv r + \tau$, one can show that the numerical elasticities of capital demand with respect to ρ_1 are given as

$$-\frac{\partial K_1}{\partial \rho_1}\frac{\rho_1}{K_1} = \frac{\widehat{\pi}\,(N_1/N_2) + (1-\beta)\left(\widehat{\pi} + \widehat{\beta}\right)}{\pi\left(\widehat{\pi} + \widehat{\beta}\right) + \widehat{\pi}\,(\pi + \beta)(N_1/N_2)},$$

$$\frac{\partial K_2}{\partial \rho_1}\frac{\rho_1}{K_2} = \frac{\widehat{\beta}\,(1 - \pi - \beta)(N_1/N_2)}{\pi\left(\widehat{\pi} + \widehat{\beta}\right) + \widehat{\pi}\,(\pi + \beta)(N_1/N_2)}.$$

Ceteris paribus, these elasticities are seen to be smaller the larger the values of π and β.

[11] See Munk (1978) for a more general analysis of optimal taxation in the presence of pure profit.

While the labor income shares and employment levels appearing in the optimal tax formula (3.14) should be easy to observe, it may be more difficult to obtain data for the pure profit shares. However, as a first approximation one might identify the fixed factor with land and use data on land values to estimate land rents. Thus it appears that the estimation of optimal relative capital income tax rates on the basis of observable variables need not be exceedingly difficult. In particular, our analysis suggests that sectors with a very high capital intensity coupled with insignificant inputs of land and natural resources are candidates for a relatively lenient tax treatment. Perhaps this observation helps to explain why so many countries have chosen to offer very favorable tax treatment of the highly capital-intensive shipping industry.[12]

3.3. Optimal Taxation and Capital Flight

More generally, many governments have tended to offer relatively low effective tax rates to industries where capital is believed to be particularly mobile across borders. Our simple model may also help us understand this tendency. To be sure, in a technical sense the model assumes that all capital is equally (perfectly) mobile, but it also implies that a given level of taxation will tend to generate a larger capital export from one sector than from the other. And when policy makers argue for a reduced tax rate on some 'mobile' activity, they are typically concerned about the risk that a 'normal' tax rate on that activity will cause a large capital flight from the country. Our model does indeed prescribe that sectors where taxation tends to cause a relatively large capital export should carry relatively low tax rates.

To illustrate this most clearly, let us return to our Cobb-Douglas example and look at the special case where both of the internationally immobile factors are also immobile between the two domestic production sectors. This case may be modelled by setting $\beta = \widehat{\beta} = 0$ in (3.12) and (3.13). From the first-order conditions in (3.2) we then find

$$\alpha K_1^{\alpha-1} = r + \tau \Longleftrightarrow K_1 = \left(\frac{r+\tau}{\alpha}\right)^{\frac{1}{\alpha-1}}, \qquad \widehat{\alpha} K_2^{\widehat{\alpha}-1} = r + t \Longleftrightarrow K_2 = \left(\frac{r+t}{\widehat{\alpha}}\right)^{\frac{1}{\widehat{\alpha}-1}}, \tag{3.15}$$

[12] To avoid misunderstanding, let me stress that I do not consider this policy to be optimal from a global perspective, but from the individual country perspective adopted here, it may make some sense.

implying

$$\varepsilon_\tau^{K_1} = \left(\frac{1}{\alpha - 1}\right)\left(\frac{\tau}{r+\tau}\right), \qquad \varepsilon_t^{K_2} = \left(\frac{1}{\widehat{\alpha} - 1}\right)\left(\frac{t}{r+t}\right), \qquad \varepsilon_t^{K_1} = \varepsilon_\tau^{K_2} = 0. \tag{3.16}$$

We see that in this scenario a higher tax rate on one sector does not channel capital into the other domestic sector. Instead, it only induces a capital export, and according to the tax elasticities in (3.16) this capital export will be larger in the more capital-intensive sector (where the alfa-coefficient is larger). According to (3.11) and (3.16) the optimal tax rule now simplifies to[13]

$$\varepsilon_\tau^{K_1} = \varepsilon_t^{K_2} \qquad \Longrightarrow \qquad \frac{\tau/(r+\tau)}{t/(r+t)} = \frac{1-\alpha}{1-\widehat{\alpha}} \tag{3.17}$$

In other words, the more capital-intensive sector (measured by the capital share of sectoral income) should carry a lower relative tax rate, since capital demand - and hence capital export - is more sensitive to taxation in this sector. In this sense our model prescribes that tax rates should be differentiated so as to minimize tax-induced capital flight.

Denoting the cost of capital in sector j by ρ_j and defining the price elasticities

$$\eta_1 \equiv \frac{dK_1}{d\rho_1}\frac{\rho_1}{K_1}, \qquad \rho_1 \equiv r+\tau, \qquad \eta_2 \equiv \frac{dK_2}{d\rho_2}\frac{\rho_2}{K_2}, \qquad \rho_2 \equiv r+t, \tag{3.18}$$

we may use (3.16) and equations (4.7) and (4.8) in the appendix to write the optimal tax rule in the alternative form

$$\frac{\tau}{r+\tau} = \frac{-\lambda/(1+\lambda)}{\eta_1}, \qquad \frac{t}{r+t} = \frac{-\lambda/(1+\lambda)}{\eta_2}, \tag{3.19}$$

where λ is the shadow price associated with the government budget constraint, measuring the marginal excess burden of taxation, i.e., the fall in national income occurring when the government raises an extra unit of revenue. Equation (3.19) is a simple *inverse elasticity rule for source-based capital taxation*, completely parallel to Ramsey's famous inverse elasticity rule for indirect taxation. The parallel is not surprising: we know that Ramsey's inverse elasticity rule applies in the special case where cross price elasticities are zero, and in our scenario with fixed domestic factors, it follows from (3.15) that the cross price elasticities of capital demand are in fact zero.

[13]Equation (3.17) may also be derived from (3.14) by setting $\beta = \widehat{\beta} = 0$ and using the definitions of π and $\widehat{\pi}$ plus the fact that $N_2 = \overline{N} - N_1$.

Of course one should not make too much of the results stated in (3.14) and (3.17), since they are based on the simplifying assumption of Cobb-Douglas production functions which implies a unitary elasticity of substitution between capital and labor in both sectors. In practice the elasticity of substitution may differ across sectors so that the exact optimal tax rates cannot be estimated from the relatively simple formula (3.14). Still, the formula does point to some observable variables such as labor and capital intensities and pure profit shares that are likely to be related to the capital demand elasticities determining the optimal capital income tax rates.

In the next section I will try to sum up the policy conclusions that would seem to follow from the analysis in the present and the previous section.

4. Summary and Conclusions

A long-standing criticism of the theory of optimal indirect taxation is that it is inoperational, since we do not have and cannot realistically obtain the information needed to differentiate indirect tax rates in accordance with the theory. In addition, there are administrative and political economy arguments against a systematic differentiation of indirect tax rates. On these grounds most practitioners and many academics argue in favor of a uniform VAT, supplemented by a limited range of Pigovian excises to correct for obvious externalities.

I certainly consider this to be a respectable position. On the other hand, the recent contributions to optimal tax theory discussed in section 2 suggest that at least some of the information that would allow a rational differentiation of indirect tax rates may be easier to collect than previously thought. Moreover, full uniformity is impossible to achieve since in practice some goods and services cannot be taxed. In this sense any indirect tax system is necessarily non-uniform, even if all *taxable* goods and services carry the same tax rate. If some tax differentiation is inescapable anyway, one could argue that policy makers might as well try to differentiate taxes in accordance with optimal tax principles, even if these principles cannot be implemented in a perfect manner.

Harberger (1990) takes an intermediate position. On the one hand, for pragmatic reasons he argues for uniform taxation of those commodities that are included in the tax net. On the other hand, he points out that it is almost always possible to include some further commodities in the tax base if policy makers are willing to incur the extra administrative cost. In that situation the important policy issue is what activities to include in the tax base and what activities to exempt. According to Harberger, it is generally inefficient to try to tax activities which are close substitutes for activities that are exempt because they are too difficult to tax. This is in line with the analysis in this paper which suggests that certain consumer services that are close substitutes for untaxed household production should be left out of the indirect tax base, or at least be taxed at concessionary rates. Indeed, the numerical simulations undertaken by Sørensen (1997) and by Piggott and Whalley (1998) indicate that the efficiency gains from reduced tax rates on such services could be substantial.

Hence my conclusion on indirect taxation is similar in spirit to that of Harberger (op.cit.). There are good pragmatic reasons for sticking to a high degree of uniformity in indirect taxation, so policy makers should deviate from uniformity only when there is a strong efficiency case for doing so. However, recent advances in the theory of optimal commodity taxation have made it easier to identify the types of commodities that would seem strong candidates for special tax treatment. In particular, certain services that are close substitutes for 'do-it-yourself' activities within the household sector should probably be exempt from tax or at least be taxed at concessionary rates. On the other hand, goods or services that require particularly large amounts of household time for their consumption should probably be taxed at relatively high rates. Neglecting these insights from optimal tax theory may cause considerable losses of economic efficiency.

In the area of direct taxation most economists seem to agree that taxes ought to be uniform or neutral across different uses of capital and labor, even if full neutrality of caital income taxation may be difficult to achieve for administrative reasons. The Diamond-Mirrlees Production Efficiency Theorem provides some theoretical rationale for neutral capital income taxation, and the work of Auerbach (1989) suggests that a closed economy can only achieve very small efficiency gains by deviating from tax neutrality when the restrictive assumptions underlying the Production Efficiency Theorem are violated. Moreover, there are strong administrative and political economy arguments in favor of neutral capital income taxation.

However, the growing international mobility of capital and the difficulties of enforcing domestic tax on foreign source income mean that even a nominally uniform capital income tax is increasingly turning into a selective input tax on capital invested in the domestic rather than the foreign economy. Just as domestic citizens may avoid tax by seeking refuge in leisure and in informal household production, they may avoid the domestic capital income tax by investing their capital abroad. In this setting we saw that a capital income tax which is uniform across all domestic production sectors will fail to minimize the unavoidable distortions caused by capital taxation, just as uniform indirect taxation will fail to minimize the distortions from commodity taxes. Indeed, I derived an intuitive Ramsey rule for capital taxation and showed that, under the popular assumption of Cobb-Douglas technologies, the optimal differentiated capital income tax rates may in principle be estimated from observable variables. More generally, the analysis suggested that the labor and land intensities of production are important determinants of the elasticities of capital demand that govern the optimal relative capital income tax rates.

Yet the same practical and political economy arguments that speak in favor of uniform commodity taxation also suggest that tax neutrality across different types of investment should remain the general norm in capital income taxation, even if the government can only tax domestic investment. In particular, governments should be careful when drawing policy conclusions from the insight that the theoretically optimal policy seeks to minimize tax-induced capital flight. If governments try to pursue this rule but do not have full information on the technological parameters influencing capital mobility, firms will have a strategic incentive to label themselves as being particularly mobile in order to qualify for favorable tax treatment.[14] However, in sectors where the tax elasticity of capital demand is known with a high degree of certainty to be either very high or very low, policy makers may want to accept some deviations from tax neutrality in order to reduce the distortionary effects of source-based capital taxation.

In summary, the practitioners' case against selective direct and indirect taxation remains strong. Hence the burden of proof should always be carried by those who argue for deviations from uniformity and neutrality, and such deviations should be accepted only in those few cases where theory and evidence clearly indicate a high welfare cost of uniform taxation. Some policy advisers fear that even the slightest concession to the adherents of

[14] For an analysis stressing this point, see Hagen, Osmundsen and Schjelderup (1998).

selective taxation will open the door to a flood of badly motivated special provisions and tax subsidies pushed by lobbyists. But in some cases the welfare loss from not allowing differential taxation may be so high that uniform taxation is politically unsustainable anyway. This paper has argued that modern optimal tax theory can help policy advisers to identify those few areas where differential commodity and capital taxation may be warranted. Indeed, by insisting that proposals for selective tax policy must be clearly defensible in terms of optimal tax criteria and empirical evidence of their quantitative importance, public finance economists may actually help to establish a bulwark against badly motivated violations of tax neutrality that only serve to promote special interests at the expense of the public interest.

APPENDIX

This appendix shows how the rules (3.11) and (3.14) for optimal source-based capital taxation may be derived. Using $N_2 = \overline{N} - N_1$, the conditions for profit maximization (3.2) and (3.3) may be condensed to

$$f_K(K_1, N_1) = r + \tau, \tag{4.1}$$

$$F_K\left(K_2, \overline{N} - N_1\right) = r + t, \tag{4.2}$$

$$f_N(K_1, N_1) = F_N\left(K_2, \overline{N} - N_1\right). \tag{4.3}$$

This system implies that K_1, K_2 and N_1 are functions of τ and t and that

$$\frac{\partial K_1}{\partial \tau} = \frac{F_{KK}F_{NN} - F_{KN}^2 + f_{NN}F_{KK}}{D} < 0, \tag{4.4}$$

$$\frac{\partial K_2}{\partial t} = \frac{f_{KK}f_{NN} - f_{KN}^2 + f_{KK}F_{NN}}{D} < 0, \tag{4.5}$$

$$\frac{\partial K_1}{\partial t} = \frac{\partial K_2}{\partial \tau} = -\frac{f_{KN}F_{KN}}{D} > 0, \tag{4.6}$$

$$D \equiv f_{KK}\left(F_{KK}F_{NN} - F_{KN}^2\right) + F_{KK}\left(f_{KK}f_{NN} - f_{KN}^2\right) < 0,$$

where the signs follow from the strict concavity of the production functions.

Assuming that $t^* = 0$, $T = \overline{T}$ and $\overline{R} \equiv R - \overline{TN} > 0$, the government's problem is to choose the capital tax rates τ and t so as to maximize national income (3.6) subject to the government budget constraint $\tau K_1 + tK_2 = \overline{R}$. Using (3.1), (3.5) and (3.6), and denoting the shadow price associated with the government budget constraint by λ, the Lagrangian corresponding to this problem may be written as

$$L = f\left(K_1(\tau,t), N_1(\tau,t)\right) + F\left(K_2(\tau,t), \overline{N} - N_1(\tau,t)\right) + r\left[\overline{K} - K_1(\tau,t) - K_2(\tau,t)\right]$$

$$+\lambda\left[\tau K_1(\tau,t) + tK_2(\tau,t) - \overline{R}\right].$$

Exploiting (4.1) through (4.3) plus (4.6), we find the first-order conditions with respect to τ and t to be

$$\frac{\tau}{K_1}\frac{\partial K_1}{\partial \tau} + \frac{t}{K_1}\frac{\partial K_1}{\partial t} = -\frac{\lambda}{1+\lambda}, \tag{4.7}$$

$$\frac{\tau}{K_2}\frac{\partial K_2}{\partial \tau} + \frac{t}{K_2}\frac{\partial K_2}{\partial t} = -\frac{\lambda}{1+\lambda}. \tag{4.8}$$

These conditions lead immediately to the optimal tax rule (3.11).

Suppose now that the production functions in (3.1) take the Cobb-Douglas form in (3.12) and (3.13) so that (4.1) through (4.3) specialize to

$$\alpha K_1^{\alpha-1} N_1^{\beta} = \rho_1, \qquad \rho_1 \equiv r+\tau, \tag{4.9}$$

$$\widehat{\alpha} K_2^{\widehat{\alpha}-1} \left(\overline{N}-N_1\right)^{\widehat{\beta}} = \rho_2, \qquad \rho_2 \equiv r+\tau, \tag{4.10}$$

$$\beta K_1^{\alpha} N_1^{\beta-1} - \widehat{\beta} K_2^{\widehat{\alpha}} \left(\overline{N}-N_1\right)^{\widehat{\beta}-1} = 0. \tag{4.11}$$

Denoting relative changes by tilde superscripts, this system may be log-linearized to give

$$(\alpha-1)\,\widetilde{K}_1 + \beta \widetilde{N}_1 = \widetilde{\rho}_1, \tag{4.12}$$

$$(\widehat{\alpha}-1)\,\widetilde{K}_2 - \widehat{\beta}\,(N_1/N_2)\,\widetilde{N}_1 = \widetilde{\rho}_2, \tag{4.13}$$

$$\alpha \widetilde{K}_1 - \widehat{\alpha}\widetilde{K}_2 + \left[\beta - 1 + \left(\widehat{\beta}-1\right)(N_1/N_2)\right]\widetilde{N}_1 = 0, \tag{4.14}$$

from which we find

$$\frac{\widetilde{K}_1}{\widetilde{\rho}_1} \equiv \left(\frac{r+\tau}{r}\right)\varepsilon_{\tau}^{K_1} = \frac{(1-\widehat{\alpha})(1-\beta) + (N_1/N_2)\left(1-\widehat{\alpha}-\widehat{\beta}\right)}{\Delta}, \tag{4.15}$$

$$\frac{\widetilde{K}_1}{\widetilde{\rho}_2} \equiv \left(\frac{r+t}{r}\right)\varepsilon_{t}^{K_1} = -\frac{\widehat{\alpha}\beta}{\Delta}, \tag{4.16}$$

$$\frac{\widetilde{K}_2}{\widetilde{\rho}_1} \equiv \left(\frac{r+\tau}{r}\right)\varepsilon_{\tau}^{K_2} = -\frac{\alpha\widehat{\beta}\,(N_1/N_2)}{\Delta}, \tag{4.17}$$

$$\frac{\widetilde{K}_2}{\widetilde{\rho}_2} \equiv \left(\frac{r+t}{r}\right)\varepsilon_{t}^{K_2} = \frac{1-\alpha-\beta+(N_1/N_2)(1-\alpha)\left(1-\widehat{\beta}\right)}{\Delta}, \tag{4.18}$$

$$\Delta \equiv (\widehat{\alpha}-1)(1-\alpha-\beta) + (\alpha-1)\left(1-\widehat{\alpha}-\widehat{\beta}\right)(N_1/N_2).$$

Using (4.15) through (4.18) plus the facts that $1-\alpha-\beta \equiv \pi$ and $1-\widehat{\alpha}-\widehat{\beta} \equiv \widehat{\pi}$, one arrives at the optimal tax rule (3.14).

REFERENCES

Atkinson, A.B. and J. Stiglitz (1972). The structure of indirect taxation and economic efficiency. *Journal of Public Economics* 1, 97-119.

Atkinson, A.B. and J. Stiglitz (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics* 6, 55-75.

Auerbach, A.J. (1989). The deadweight loss from 'non-neutral' capital income taxation. *Journal of Public Economics* 40, 1-36.

Auerbach, A.J. (2006). The choice between income and consumption taxes: a primer. NBER Working Paper 12307.

Auerbach, A.J. and J.R. Hines (2002). Taxation and economic efficiency. In *Handbook of Public Economics*, vol. 3, edited by A.J. Auerbach and M.S. Feldstein. Amsterdam: North-Holland. 1347-1421.

Becker, G.S. (1965). A theory of the allocation of time. *Economic Journal* 75, 493-517.

Bovenberg, A.L. and L. Goulder (1996). Optimal environmental taxation in the presence of other taxes: general equilibrium analyses. *American Economic Review* 86, 985-1000.

Christiansen, V. (1984). Which commodity taxes should supplement the income tax? *Journal of Public Economics* 24, 195-220.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54, 607-622.

Corlett, W.J. and D.C. Hague (1953). Complementarity and the excess burden of taxation. *Review of Economic Studies* 21, 21-30.

Diamond, P. og J. Mirrlees (1971). Optimal taxation and public production (I). *American Economic Review* 61, 8-27.

Hagen, K.P., P. Osmundsen and G. Schjelderup (1998). Internationally mobile firms and tax policy. *Journal of International Economics* 45, 77-96.

Harberger, A.C. (1962). The incidence of the corporation tax. *Journal of Political Economy* 70, 215-240.

Harberger, A.C. (1990). The uniform-tax controversy. In: Tanzi, V. (ed.), *Public Finance, Trade, and Development.* Wayne State University Press, Detroit, MI (proceedings of the 44th IIPF congress). 3-17.

Huizinga, H. and S.B. Nielsen (1997). Capital income and profits taxation with foreign ownership of firms. *Journal of International Economics* 42, 149-165.

Judd, K. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28, 59-83.

Keen, M. and J. Ligthart (2006). Incentives and information exchange in international taxation. *International Tax and Public Finance* 13, 163-180.

Kleven, H.J. (2004). Optimum taxation and the allocation of time. *Journal of Public Economics* 88, 545-557.

Kleven, H.J., W. Richter and P.B. Sørensen (2000). Optimal taxation with household production. *Oxford Economic Papers* 52, 584-594.

Munk, K. (1978). Optimal taxation and pure profit. *Scandinavian Journal of Economics* 80, 1-19.

Musgrave, P.B. (1969). *United States Taxation of Foreign Investment Income: Issues and Arguments*. Cambridge, Mass.: Harvard Law School, International Tax Program.

Piggott, J. and J. Whalley (1998). VAT base broadening, self supply, and the informal sector. NBER Working Paper 6349.

Pigou, A.C. (1920). *The Economics of Welfare*. Macmillan.

Ramsey, F. (1927). A contribution to the theory of taxation. *Economic Journal* 37, 47-61.

Razin, A. and E. Sadka (1991). International tax competition and gains from tax harmonization. *Economics Letters* 37, 69-76.

Sadka, E. (1977). A theorem on uniform commodity taxation. *Journal of Public Economics* 7, 387-391.

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics* 83, 217-230.

Salanié, B. (2003). *The Economics of Taxation*. MIT Press, Cambridge Massachusetts.

Sandmo, A. (1974). A note on the structure of optimal taxation. *American Economic Review* 64, 701-706.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *Swedish Journal of Economics* 77, 86-98.

Sandmo, A. (1976). Optimal taxation: an introduction to the literature. *Journal of*

Public Economics 6, 37-54.

Sandmo, A. (1990). Tax distortions and household production. *Oxford Economic Papers* 42, 78-90.

Sandmo, A. (2000). *The Public Economics of the Environment.* Oxford University Press.

Stiglitz, J. og P. Dasgupta (1971). Differential taxation, public goods, and economic efficiency. *Review of Economic Studies* 38, 151-174.

Sørensen, P.B. (1997). Public finance solutions to the European unemployment problem? *Economic Policy* 25, 223-264.

Sørensen, P.B. (2006). Can capital income taxes survive? And should they? Invited lecture at The CESifo Venice Summer Institute Workshop on *The Future of Capital Income Taxation*, July 17-18, 2006.