I. Introduction: Choosing Among Tax Systems

There are three main approaches to analyzing the tax systems observed in countries at a point in time and changes in them through time. One is simply historical. It recounts the tax systems in place, and cases in which the government (legislature, prime minister, king, or dictator) adopted a new tax system or significantly revised the existing system. Many such accounts attempt simply to tell the story of taxation, without much attention given to the process through which taxes come to be adopted or the interests advanced or harmed by them. There are relatively few such treatments, in part because few historians or readers are very interested in such accounts. For example, in the West, tax revenues in the early nineteenth century consisted mainly of property taxes, excise taxes, and tariffs. Income taxes were added in many European tax revenues during the late nineteenth century. The federal income tax in the United States was added in 1913. Sales and value were added taxes in the twentieth century. Germany and France introduced sales (consumption) taxes during WWI. The first sales tax in the United States was introduced by West Virginia in 1921. The first national VAT tax was introduced in France in 1958 (after experimenting with such a tax in their Ivory Coast colony for a few years).

A second approach is to focus on taxes that "should" be chosen because they have desirable economic and/or normative properties. Among economists, this approach is called optimal tax theory. It attempts to identify the "best" tax system, based on its conformity to normative properties and economic impacts. Optimal tax theory is usually undertaken from the perspective of utilitarian philosophy, which reflects the utilitarian roots of neoclassical economics from the nineteenth century. For example, a utilitarian central planner may be assumed to adopt a tax system—and the analyst then examines the type of system adopted, given various assumptions about the nature of the economic impacts of the tax and aggregate utility function. Most, but not all of these types of analyses assume that the pattern of expenditures is predetermined, and the issue is simply how to raise the money to pay for those expenditures. Several ideas about the properties of ideal tax systems have come out of this approach. For example, broad based taxes are better than narrow tax systems or tax systems riddled with loopholes because they have lower average tax burdens. Another is that neutral taxes have advantages over less neutral ones, except where externalities are to be addressed through the tax

system. Utilitarian economists also tended to favor progressive over flat or regressive taxes and have done so since shortly after the marginal revolution. Given diminishing marginal utility, wealthy persons tend to lose less utility when they give up their last dollar of income to tax authorities than do poorer persons. Thus, to equalize burdens in utility terms one should collect more money from the relatively wealth than the relatively poor.

The third approach is to construct models of how tax systems would be designed under various forms of government. An authoritarian might, for example, attempt to maximize net revenues from the tax systems as argued by Mancur Olson (1993).

A democracy would, in principle, tend to adopt the tax system preferred by the median voter—if one exists. Among voters, however, taxation—unlike many forms of expenditures—is essentially a zero-sum game. "I" would rather that "you" pay more taxes so that "I" can pay less. And you would rather than "I" pay more taxes, so that "you" can pay less. This tends to make tax systems unstable for reasons developed below. Identifying common interests that voters might have regarding taxation is thus difficult.

However, it turns out that voters generally share interests in tax systems under which tax burdens are relatively easy to determine and in which taxes are generally stable from year to year. Consensus about the details of systems of taxation appear to require normative support of one kind or another. Optimal tax theory may play a role in this process insofar as the normative theories proposed or conclusions reached are internalized by voters. When they are, such normative theories may contribute to a tax system's stability by bounding the domain in which cycles occur.

One significant for problem confronted by democratic regimes is that a median voter does not always exist with respect to tax systems. This had led some public finance theorists to adopt models of voter or politician decision making that depart from the usual rational choice models of neoclassical economics. The most widely used alternative to the median voter model is the stochastic voter model. That model, voters are assumed to be more or less rational in that they are more likely to best candidate or policy than others, but also may vote for candidates with less attractive policies. Such models have, for example, been used by Hettich and Winer (1984, 2005, 2018) to analyze complex tax systems.

The outcomes of stochastic voter models tend to be weighted averages of voter preferences rather than determined by a single individual—where the weights vary with each voter's propensity to change his or her votes in response to small changes in public policy (See Coughlin 1992, Nitzan and Coughlin 1981), These models do not require as many assumptions or as many unrealistic ones as deterministic choice models to generate sharp predictions, although cycling problems are not entirely avoided by that methodology (see Kirchgässner 2000). However, they do require pulling back from the normal rational choice models used in economics and public choice.

Nonetheless, empirical work based on median voter models with preexisting tax systems and products are commonplace, and they tend to account for variations in expenditures in a manner that is consistent with those models. It is possible, that a richer model of median voter interests—one that includes the possibility of normative interests will prove sufficient to create median voter models of tax systems.

This chapter first reviews some of the basic conclusions drawn from optimal tax theory and then examines how such theories, if internalized by a sufficient number of voters or legislators may influence the tax systems that we actually observe.[1]

II. Normative Theories of Taxation—"Optimal" Tax Theory

Although theories of how to rank societies or public policies are arguably a subject for political philosophers and theologians rather than economists, economic analysis is often motivated by efforts to improve a society or public policy. And, it turns out that many of the "intuitions" that economists hold about how policies may do so have utilitarian foundations. Cost=Benefit analysis can be thought of as a dollarized form of utilitarian accounting that attempts to determine whether net benefits or aggregate utility is increased or diminished by a particular policy.

Normative analysis attempts to determine the nature of what might be called "the good

[1] I have, for example, worked on several aspects of the U.S. fiscal package and of the services provided by OECD member states. See for example Congleton and Shughart (1990) [on US social security and Medicare], Congleton and Bennett (1995) [on US highway expenditures], Congleton (2001) [on growth of US governmental spending], Congleton and Bose (2010) [on the emergence of Western welfare states], or Congleton, Batinti, Pietrantonio (2017) [on Western healthcare systems].

society," and the best means for moving a given society towards such a society. For utilitarians, such a society would maximize aggregate utility, but other normative theories would imply other ideas about the nature of a good society and the best way to achieve one given our present circumstances. With respect to tax policies, the ethical, moral, or normative question is: how should we acquire the resources required to produce government services?

Such normative considerations may or may not actually affect the manner in which tax systems are actually created and implemented. An authoritarian government could, for example, simply take what it wanted from its citizens using threats of coercive power to do so—its police force and army. It turns out, however, that for reasons doubtless similar to those described in Olson's model of stationary bandits, that authoritarian regimes normally (although not always) use rule-based confiscation (taxes) rather than willy-nilly taking to acquire their resources. Such rule-based systems of "takings" reduce the tax base by less than that associated with simply taking the desired assets of persons within the territory governed. Few people would bother to acquire "surpluses" or "assets" if they expected them to be simply taken by the most powerful organizations in their neighborhood. If the aim of a ruler, ruling committee, or political party is simple to maximize long term "extraction," as in the Olson model, then maximizing the present value of tax receipts is the aim of a "good" tax law.

If, on the other hand, the aim is to pay for services that a citizenry believes it would be useful for a government to provide—as in a "productive" or "social contract" based government—then the tax laws adopted should be acceptable to all or most citizens within the territory governed. In the latter cases, tax laws are likely to include both pragmatic and normative aims.

In either case, taxes will share a few properties. A tax will tend to be relatively easy to collect and tend to have a smaller dead weight loss than other tax systems that could be imposed—assuming that the government of interest is reasonably stable. It may also be consistent with a community's theory of good policies and fairness to reduce the extent to which taxpayers try to avoid paying taxes.

The models of governments sketched out in previous chapters have all stressed the pragmatic interests of voters and rulers. The models developed by "optimal tax theorists"

stress the ethical, moral, or normative dimension of taxation. Latter in this chapter, we take up possible normative interests that voters, legislators, and their advisor might have in addition to their practical interests. Both practical and normative interests may influence preferences over tax systems and government services.

III. Utilitarian Distributions of Tax Burden and Public Services

A wide range of normative theories can be used as guides to taxation. For example, the Pareto criteria evaluates the merits of fiscal systems by determining whether a particular system can be improved on or not, where "improve" means to make at least one person better off without making anyone worse off. If that is possible, then the present system is not "optimal" or "ideal." If such changes are not possible, then a tax or fiscal system is Pareto optimal or Pareto efficient. Whether a policy is ideal or not in this sense depends on its consequences of the proposed changes on the individuals affected by those changes. If it makes at least one person better off and no one worse off, it is an improvement—a Pareto superior move. A policy or outcome is Pareto optimal if no Pareto superior moves are possible, which is to say if no policies can generate outcomes that make at least one person better off and no one worse off.

Alternatively, a policy may be judged optimal from the point of view of some single dimensioned "optimand" such as social welfare, aggregate utility, or per capita RGNP. A policy that maximizes such an optimand is optimal or good from the perspective of its associated norm and normative theory.

Utilitarians of the late eighteenth and nineteenth centuries believed that good policies required taking account of the effects of every policy on everyone within a country. And, because everyone was ultimately interested in happiness or utility, they concluded that maximizing the sum of everyone's happiness or utility should be the goal of every government and every government policy.

The notion of utility shed light on the behavior of individuals and of policy makers and also provided a rough gauge of the relative merits of different policies and political institution. And it turned out that many of the most important economists of the nineteenth century were utilitarians. Thus, utilitarian ideas about human nature and the nature of a good society had a major impact on both positive economics and normative economics, which is also known as

welfare economics. According to this approach the best public policies maximize aggregate utility both in the present and on into the future.

Conditions that characterize the ideal utilitarian outcome can be easily developed using a bit of calculus. For example, in a setting in which a single-dimensioned pure public good (G) is financed by a proportional tax on income, the ideal level of G is that which maximize the sum of everyone's utility, given the assumed tax system. The necessary tax for this fiscal system is $tY=c(G)$, where Y is national income, t is the proportional tax rate, and c(G) is the cost of the government service. The ideal tax rate under that system would be that required to fund the service, $t^* = c(g^*)/Y$, where t^* and g^* jointly maximize aggregate utility.

If each person in the community or territory of interest has a utility function of the form $U_i= u(C_i, G)$ where $C_i = (1-t)Y_i$, then aggregate utility can be written as, $W = \sum U_i$.

After a bit of algebra and substitution, both the tax rate and after-tax income can be written as a function of G. This allows the Benthamite aggregate utility function to be written as:

$$W = \sum U_i(1-C(G)/Y)Y_i, G) \qquad (1)$$

The conditions for maximizing aggregate utility can be found by differentiating equation 1 with respect to G and setting the result equal to zero (assuming that W is strictly concave).The welfare maximizing output of the public service under this tax system requires service level G to satisfy:

$$W_G = \sum\{U_{iC}(-C_G/Y)Y_i + U_{iG}\} = 0 \qquad (2a)$$

Or, distinguishing between the marginal benefits and cost of the government serve, this can be rewritten as

$$W_G = \sum U_{iG} - \sum U_{iC}(C_G/Y)Y_i = 0 \qquad (2b)$$

Equation 2b can be interpreted as characterizing the ideal level of government service G as that which the sum of the marginal benefits from the government services(the sum of the terms, U_{iG}) equal to the sum of the marginal opportunity cost of financing that service (the sum of the $U_{iC}(-C_G/Y)Y_i$ terms). Notice that the latter is the sum of every individual's marginal opportunity cost for the government service, where the opportunity cost is measured in terms of lost utility from private consumption and rate at which private consumption is lost as

G increases. The marginal opportunity cost terms each combine subjective assessments (marginal utility) and objective terms (marginal cost shares), which implies that this calculation is not entirely objective although it is partially so. (Critics of the utilitarian approach argue that the subject part is difficult if not impossible to quantify.)

The social welfare functions used in post-WWII papers and texts are "generalizations" of the Bentham characterization of aggregate utility (as a sum) used above. They allow for different functional forms such as multiplicative ones (sometimes called a Paretian social welfare function) and also for weighting individuals differently. The latter will include many forms of a social welfare functions that would be rejected by contemporary norms such as treating heroes and villains differently, or nobles and commoners differently, or slave owners and slaves differently. (Most of the possibilities are rather nasty "generalization" that conflict with both utilitarian norms and contemporary normative notions such of equality before the law.) The simpler Benthamite aggregate utility function is far more normatively attractive than the more general and abstract social welfare functions used in most present optimal tax theory research.[2]

Another weakness of most optimal tax theory papers is that decisions are often assumed to be made by a "benevolent social planner" or "benevolent dictator" who uses a social welfare function to choose public policies.[3]

Samuelson's Characrization of the Optimal Provision of a Pure Public Good

The essential mathematics of the Samuelsonian characterization of the Pareto optimal collective provision of a pure public good is developed in this subsection. This is a generalization of the model worked out above in that no specific tax system is assumed.

The first step is to develop some notation: Let G be the level of a pure public good, let Xi be the level of a pure private good received by individual *i*, let Ui = u(G, Xi) be the utility of individual *i* associated with a particular combination of the public good G and private good Xi

---

[2] For background on social welfare functions, see Arrow (1950), Bergson (1954), Harsanyi (1975), Samuelson (1977), Bergstrom (1993), and/or Florio (2014).

[3] This is evidently an extension of Plato's idea of the philosopher king being the best form of government, but it is of course actually a form of dictatorship unless all the persons in the community agree about the exact nature of the best social welfare function.

received by *i*. Second, we need some macro-choice notation. Let W be a social welfare function and let T(G,X) = 0 be the technological frontier of combinations of the public good and private goods, with X = ΣXi. The constraint function is in its "zero form." Assume that there are N persons in the society of interest.

The problem of maximizing social welfare can be undertaken with the approach developed by Lagrange, e.g., first create a Lagrange function, then differentiate with respect to the control variables and set the result equal to zero:

$$\mathcal{L} \max = W(U_1, U_2, U_3 .... U_N) - \lambda(T(G,X)) \quad (3)$$

Differentiating the Lagrangian with respect to G, X_1, X_2, X_3 .... X_N, and λ yields the first order conditions for the combination of services and private incomes that maximizing social welfare:

$$\Sigma W_{Ui} U i_G - \lambda T_G = 0 \quad (4)$$

$W_{Ui} U i_X - \lambda T_X = 0$ for all i = 1 ... N (This represents N first order equations)

$T(G,X) = 0$

After obtaining the Lagrangian first order conditions, the next step is to manipulate the first order conditions into a form that is both economically interesting and useful. Samuelson uses a rather clever series of steps to do so.

First, shift the lambda terms to the right and then divide the first and second first order condition to eliminate the lambda.

$$[\Sigma W_{Ui} U i_G] / W_{Uj} U j_X = T_G / T_X \quad (5)$$

(I have used the j-th of the private good first order conditions to avoid confusion with "i" the counter for the summation in the public goods terms in the numerator)

Since the denominator does not change with "i" it can be brought inside the brackets--because it is essentially a constant as far as this fraction is concerned.

$$\Sigma [W_{Ui} U i_G / W_{Uj} U j_X] = T_G / T_X$$

Now note that the first order conditions for the private goods derivatives imply that the marginal social welfare generated by the private goods allocated to each individual are equal to one another at the margin:

$$W_{Ui} U i_X = W_{Uj} U j_X \quad (6)$$

This condition holds for all "i" and "j" (for every person's private good). This equivalence means that you can rewrite the equation under part b as:

$$\Sigma \left[W_{Ui} U_{iG} / W_{Ui} U_{ix}\right] = T_G / T_X \qquad (7)$$

Note that the denominator now has a counter that moves with the sum. (This is the clever part of the derivation. The rest is pretty straight forward.) We have substituted the various "I-terms" for the "j" term that we started with.

This allows us to simplify quite a bit—the marginal weights W_{Ui} in the numerator and denominator are the same and can be divided out of each. This implies that at G*:

$$\Sigma \left[U_{iG} / U_{ix}\right] = T_G / T_X \qquad (8)$$

The ideal level of a pure public good will set the **sum of the marginal rates of substitution between the private and public good equal to the technological rate of transformation** between them. (In benefit/cost terms, G* occurs at a point where the sum of the marginal benefits equals the marginal cost of the public good in terms of reductions in the private good.) Surprisingly, **the optimal level of a pure public good is completely independent of the social welfare function** used!

However, the distribution of the private good will vary with the specific for of the social welfare function assumed. Note also that this a pure "benevolent central planner" model in which the imaginary planner can make both the supply of pure public good decisions and the has complete control over the allocation of private income or consumption decision without input from the individuals in the society!

IV. The Contractarian Approach to Optimal Taxation

Exceptions to mainstream reliance on utilitarian ideas and social welfare functions include the Pareto criteria (although Pareto himself was a utilitarian) and the normative contractarian theories developed by John Rawls (1971) and James Buchanan (1962 with Tullock, 1988, with Brennan).

Contemporary **contractarians** such as John Rawls and James Buchanan regard consensus to be the best indication of whether a procedure for making policy choices or a specific policy change is an improvement or not. If everyone agrees that social state Z' is better than Z", then it is—at least as far as contractarians are concerned.

Agreement is considered a better indicator than aggregate utility, because aggregate utility cannot be reliably measured and because good policies should—according to this approach—make everyone better off, rather than simply assure that the "winners" gain more than the "losers" lose as in utilitarian analysis.

Consensus requires each individual to assess his or her expected net benefits from the consequences of policy G' and G" and then to vote in favor if his or her expected net benefits (adjusted for risk) are greater than zero. The problem confronted by contractarians is that both honest (non-strategic) voting and unanimous agreement are rarely forthcoming in "real world" circumstances. Individuals often have different interests and so disagree about the best policies. And, when voting, they may vote strategically. Even if a proposal makes them better off, they may vote against it in the hope of getting a policy that is closer to their ideal than the one being proposed.

To be fair to the contractarians, they do not generally favor consensus-based politics on every issue but tend to argue that citizens might agree to procedures for selecting policies (Buchanan) or principles for ranking institutions and policies (Rawls). They believe this to be the case because it is more difficult for individuals to assess their own narrow benefits from long-term institutional procedures and principles than it is on day-to-day policies.

This uncertainty is sometimes referred to as "the veil." For Buchanan, it is the real uncertainly associated with commitments to standing procedures that tends to generate agreement (the veil of uncertainty). Would you prefer to use majority rule for every choice, or super majority or unanimity, or perhaps some combination of all three: unanimity for the most important, super majorities for other important issues, and majority rule for routine decisions, etc.?

For Rawls, the "veil" is a method of thinking about principles that tends to produce consensus about the best principles for designing governing institutions, the Rawlsian "veil of ignorance." From behind the veil, Rawls assume that one disassociates one's real self from the decision and instead imagines what would be best if he or she did not know what role or position he or she would occupy in the society chosen via a principle of justice. This allows alternative principles of justice to be assessed. Each individual imagines what it would be like to be in

all the various positions in society and think about what principles (or what institutions) they would like public policies to reflect in that situation.

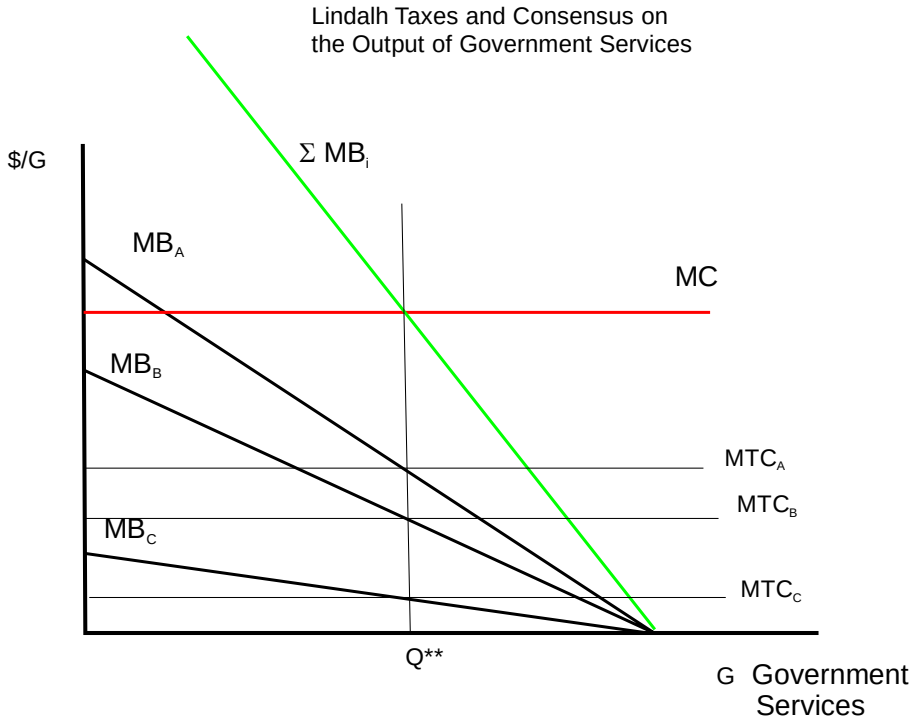
For example, if you didn't know which person you would be in the following 4 societies of 3 people, which society would you prefer? (5, 5, 5), or (8, 8, 5), or (12, 7, 4), or (18, 5, 2) if there was an equal chance that you could/would be each of the individuals?

Consensus-Based Optimal Taxation and Pure Public Goods: Lindalh Taxes

It turns out that there is a special case of the Samuelsonian social welfare maximizing provision of a pure public good that tends to generate unanimous agreement. It was an idea developed somewhat before Samuelson's famous paper. It is attributed to Erik Lindahl, who was interested in tax systems that would tend toward unanimous approval and/or minimize resistance.

Recall that in its geometric form, the Samuelson Conditions required (i) producing the service at the point (G**)where the sum of the marginal benefits realized by all individuals equals the marginal production cost of the good, (ii) the sum of the individual tax payments to add up to the cost of producing G** units of the pure public good, and (iii) that the sum of the marginal tax payments equal the marginal cost of producing the service. To that list, the Lindahl tax system adds the provision that each person's marginal tax cost should equal's his or her marginal benefit at output G**. It is for this reason that Lindahl taxes are sometime called benefit taxes.

The figure below illustrates such a tax system for three persons or three homogeneous groups of taxpayer citizens. Al, Bob, and Cathy each have different marginal benefit curves, which, as drawn, may be simply because of differences in income, although their tastes for the service may also differ. The Pareto optimal level of the public service is G** is where the sum of those marginal benefit curves (the social marginal benefit curve) crosses the marginal cost of producing the service line (MC). The Lindalh taxes shown are one of many possibilities for which the individual marginal tax costs (MTC) equal their marginal benefits at G**. Notice that, given their tax costs, each person prefers G** to every other level of service. Thus, a Lindahl tax system can generate unanimous agreement about government service levels.



That unanimity is even conceptually possible is a bit surprising given the differences in interests indicated by their marginal benefit curves. In practice, the challenge would be estimating each person's marginal benefit curve.[4]

V. Narrower Tax Norms: Minimizing Excess Burden

These three approaches—the utilitarian, the contractarian, and the Paretian are the most widely used normative foundations for economic analysis, with the utilitarian one as interpreted by Pigou being by far the most common one. However, familiarity with these approaches is also useful for economists as individuals—which do you find most plausible as a method for identifying good policies?

This section focuses on a couple of implications of the utilitarian approach to tax systems that are somewhat narrower than those of the fiscal systems characterized by Samuelson and Lindalh. The analytics rely upon the social-net benefit maximizing norm sketched out by Pigou and now routinely used in benefit-cost analysis of public policies and also in geometric analysis within public economics. A deeper study of normative theories for choosing among

[4] A method for eliciting honest revelations of voter marginal benefit curves was worked out by Clarke (1971) and discussed in a simplified form by Tideman and Tullock (1977). See Kawagoe and Mori (2001) for an experimental test of the Clarke tax mechanism.

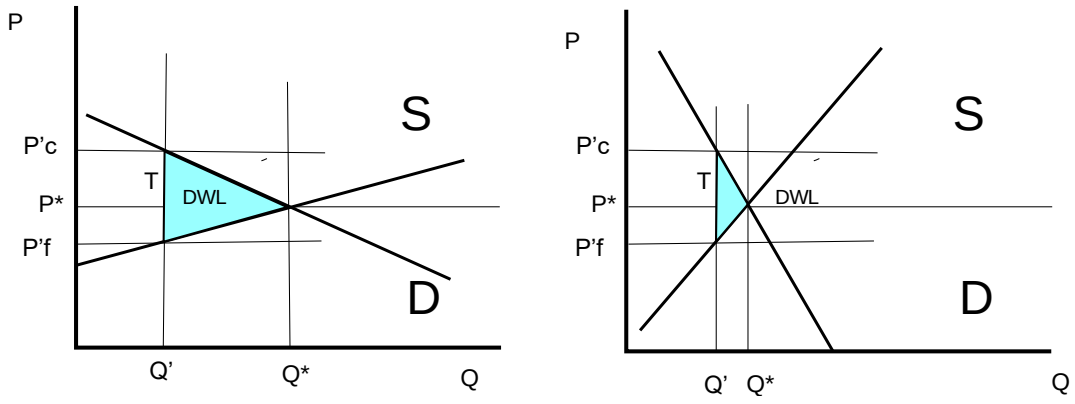
public policies is, of course, of interest, but beyond the scope of this course.

Designing Tax Systems to Minimize Excess Burden

The geometric approach and their calculus-based counter parts tend to assume that expenditures are fixed and that the issue is how best to raise the money required to fund those expenditures. We have already seen that this assumption is inconsistent with the positive theory of taxation in that both voters and dictators take simultaneously account of expenditures and taxes. Neither can choose an expenditure independently of the tax system that will be used to fund them, nor a tax rate without considering the benefits that will be associated with the expenditures to be funded.

Nonetheless, it is sometimes a useful simplification to imagine that expenditures are chosen before taxes—but only if it truly simplifies one's analysis of taxation in a manner that makes some conclusions sharper than they would otherwise have been—without obviously conflicting with the positive theory of taxation. For example, the choice of a tax system may be separate from the choice of tax rates. It may be possible to "easily" rank tax systems in a manner that would be consistent with median voter or authoritarian goals, without being easy to determine the best tax rate or level of government service. For example, some taxes may be "obviously" inferior to others because they have a larger burden. An excise tax on a market with relatively price sensitive (price elastic) demand and supply curves will have a larger excess burden than the same tax on a market with relatively price insensitive (price inelastic) demand and supply curves. The greater the excess burden associated with a tax system, the lower net benefits from non-governmental services tend to be after revenues are raised, other things being equal. This is the rationale for Ramsey (1927) tax systems—that minimize the excess burden of every level of revenue that might be sought.

Different DWLs for identical tax revenues and rates



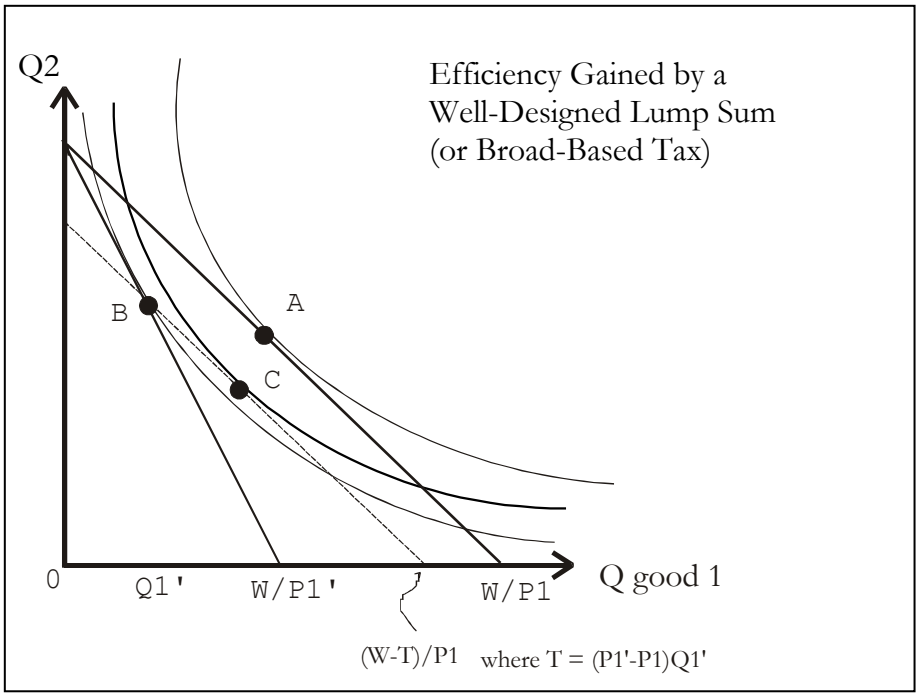
The same argument applies to subsidies. That is to say, if one is interested in minimizing deadweight losses, one should focus subsidies on markets with relatively little price sensitivity. Moreover, when earmarking an excise tax for use as a subsidy, both markets should be ones with relatively little price sensitivity (steep slopes).

Notice, however, that these deadweight-loss arguments do not address the question of why one has a tax or undertakes a subsidy. These evidently are based on something other than minimizing deadweight losses or neither excise taxes nor targeted subsidies would ever be adopted. The positive theory of taxation simply says that taxes and subsidies are targeted because moderate voters prefer them to be that way, and their excess burden(s) is fully taken into account by voters, except insofar as voters suffer from fiscal illusion(s) associated with their rational ignorance. In the latter cases, economists might usefully point out that voters would be better off with different tax and/or subsidy systems—or simply note that they prefer other tax and subsidy systems themselves. The issue for positive tax theory is what "drives" voter preferences over policies rather than normatively assessing voter preferences or public policies.

In general economists tend to favor "neutral" general taxes over targeted taxes because such taxes tend to have lower excess burdens than other tax systems. That is to say their total burdens associated with any given level of revenue (revenue raised plus excess burden) tends to be smaller than that of targeted taxes (with one exception, that we'll discuss later).

The indifference curve diagram below illustrates the geometric case for broad-based neutral taxes and/or lump sum taxes. It assumes that there are two goods that consumers are

interested in and that initially only one of them is taxed, as with an excise tax. In the pre-tax setting the individual of interest, "Al", purchases bundle "A". The effect of an excise tax on good 1 is to increase the price of the taxed good from P* to Pc . This increase in price affects the shape of each consumer's budget set. It rotates the budget constraint from the untaxed end of the budget constraint and generates a new budget constraint that lies inside the original one at all points where the consumer purchases positive quantities of the taxed good. In the case drawn, the new higher price causes the consumer to purchase bundle B instead of A. (Indeed, A is no longer feasible.)



The lump sum or general tax that raises the same revenue is characterized by the budget line parallel to the original pre-tax budget line and passing through point B. Note that under that tax scheme, the consumer can reach a higher indifference curve and chooses bundle C rather than bundle B. The difference between the utility realized with bundle C over that of bundle B the excess burden of a non-neutral tax on consumers.[5]

As in the demand and supply diagrams, much of the excess burden (deadweight loss) is

[5] This can be demonstrated by showing that a lump sum tax equal in amount to the amount this consumer pays under the excise tax, generates a budget line parallel to the first passing through bundle B. This algebraic exercise is left to students to work through.

a consequence of reduction in purchases of the taxed good, particularly that part which was generated by the "relative price" effect of the excise tax. A neutral tax has no direct effect on relative prices and so has a smaller effect on sales of goods, one that operates through the income effect alone rather than through both an income effect and substitution effect. [However, it bears noting that no tax or tax system can be completely neutral, because all taxes that are "widely known" to exist will affect locational choices of firms and consumers among towns, counties, states, and nations.]

VI. Regulating Externalities with Taxes: Pigovian Taxation, and the Double Dividend

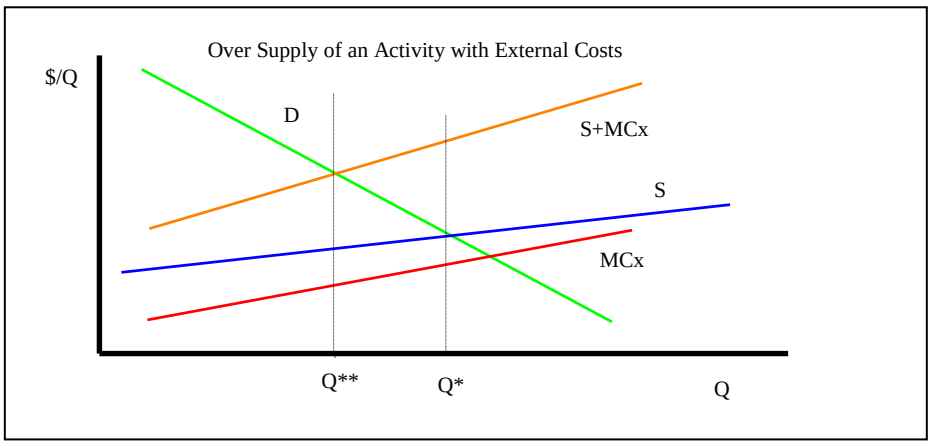
Another tax system that is consistent with the Ramsay norm of minimizing the excess burden of taxation is the Pigovian tax. A Pigovian tax is designed to solve externality problems and thus Pigovian taxes increase social net benefits rather than reducing them (assuming that the tax revenue generates at least an equal amount of benefits dollar per dollar). A Pigovian tax is designed to change behavior, in contrast to most excise taxes which are adopted simply to generate tax revenue. Pigovian taxes are sometimes argued to have a "double dividend" they solve externality problems and produce revenue in an efficient manner.

The geometry of externalities and externality problems is straight forward. In a supply and demand (market) diagram, we introduce a new curve that represents the **external** marginal costs (or marginal benefits) of the activity of interest. The predicted market outcome—Q* where the Demand and Supply curve cross—is not affected by the existence of the new marginal external cost curve, because both firms and consumers are assumed to ignore the externality generated. (Note that this positive prediction plays an important role in the entire exercise and is an assumption that can be tested.)

To find out whether an externality generating activity or output is over or under supplied, we find the social marginal benefit and marginal cost curves and use them to characterize the social net benefit maximizing activity level (output, Q**). To find the SMB and SMC curves, recall that the Demand curve is approximately the same as the marginal benefits received by consumers and the supply curve is approximately the industry's marginal cost. To these we add the external marginal benefits and/or external marginal costs to find the social marginal benefit and social marginal costs curves--now taking account of the spillovers.

Because an externality generating activity generates benefits or costs for a wide range of people simultaneously, the social marginal benefit and marginal cost curves for such activities are "vertical" sums of the relevant consumer, firm, and spillover MB and MC curves. The level of the activity that maximizes social net benefits is generally found where the social marginal benefit of the activity equals its social marginal cost curve. **If Q* does not equal Q**, there is an externality problem.**

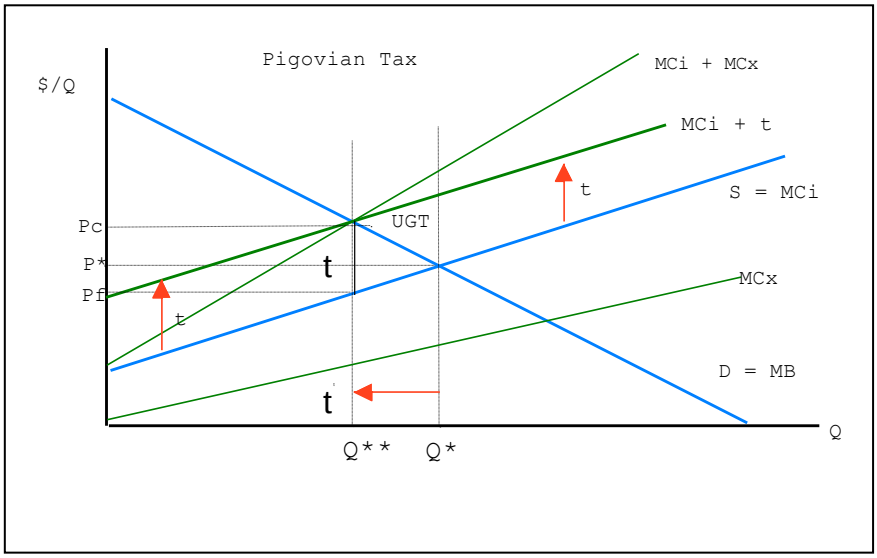
Illustration: In the figure below, the market supply and demand cross at Q*, but the SMB and SMC curves (here D and S+MCx) cross at Q**. Since Q* is not equal to Q** there is an externality problem.



The inefficiency (market failure) conclusion of this diagram can be reached using several normative theories. For example, (i) the activity level chosen fails to maximize social net benefits, then there is an externality (or public good) problem. (ii) The activity fails to realize all potential gains to trade and so there are Pareto superior moves possible. (iii) From behind a veil of ignorance, the community would (or may) agree to implement an institution that generates Q** rather than Q*.

A Pigovian tax attempts to change incentives at the margin by imposing a tax (or subsidy) on the activity that generates the externality. The ideal Pigovian tax equals the marginal external cost of the externality generating activity at Q**. Such a tax (or subsidy) is said to internalize the externality, because it makes the externality producer bear the full cost of his actions (at Q**). The figure below illustrates the effect of a Pigovian tax on an externality problem similar to that above. Without a Pigovian tax, there are unrealized gains to trade (see triangle UGT) at Q*, between the firm and those affected by the externality. The external cost at

Q** is also the vertical distance from supply curve (MCi) to the demand curve at Q**. Thus, a tax equal to the marginal spillover cost at Q** induces the market equilibrium to shift from Q* to Q**.



The extent of the revenues generated vary with the size of the externality and markets of interest. Pigovian taxes, such as an optimal carbon tax, may yield substantial revenues although this is not their main purpose. Their main purpose is to change behavior.

Notice that Pigovian taxes do not have a deadweight loss because social net benefits increase, rather than decrease, when a Pigovian tax is imposed, because an externality problem is solved. However, any tax higher than equal to the marginal external cost at Q** will generate an excess burden. (See Goulder (1995) for a useful overview of the literature on the double dividend. See also Blovenberg (1999), and Parry and Bento (2000).)

The most important of Pigovian taxes under discussion during the past decade or two is the **carbon tax**, which would generate levels of revenue similar or greater than that of the Corporate Income tax.

VII. The Characteristics of Ideal Tax Systems for Democracies

An electoral based analysis of "optimal" tax systems starts from the perspective of voters and focuses on interests that all or most voters tend to share about tax systems, rather with a hypothetical "benevolent central planner" or philosopher king. Although the allocation of tax burdens is inherently conflictual—a zero sum game, voters do have a few shared interests.

There are informational that voters share and there are normative interests that voters may—but do not necessarily—share. (This section of chapter 5 is based on Congleton (*Kyklos* 2023), where a longer, more complete analysis is developed.)

Shared Interests with respect to Delegation

If voters had complete trust in government officials (whether elected or not), they might choose to delegate complete authority over taxation to a "tax czar" or committee of "tax czars" that serve lifetime appointments. Such a system would largely isolate tax policy from political pressures—except at the moment when such tax law writers are selected. In that case, the relevant choice would be the particular philosophical and economic schools from which the best such persons would be selected. Should they be Lockeans with Smithian Intuitions, Utilitarian with Marshallian economic dispositions, or Rawlsians with a Ramsay approach to Tax theory? If such political agents could be fully trusted to adopt laws with respect to particular philosophical and economic theories of the good society, then the main choice would be which approaches are best. Perhaps a committee with a mixture would be best.

Voters could benefit from complete delegation in such cases, because most are completely ignorant of deep philosophical issues involved in ranking types of social outcomes and also equally deep and complex social science issues associated with their implementation and long run effects on society. Experts might be able to navigate such issues with greater success than voters. However, if such experts are not entirely devoted to their philosophical perspectives, they might abuse their authority by adopting policies that favor friends and family directly or indirectly through effects on industries in which they might have interests or employment. Or, they may simply devise tax systems that maximize their own net advantages by taking bribes or making extortive threats on industries subject to their tax policies. Moreover, mistakes in selecting such persons would be almost impossible to correct.

If trust is less than complete, then voters will choose to retain at least some authority over taxation. In that case, voters might delegate elected government officials the authority to adopt tax policies or appoint tax experts whose policies would be subject to their veto or supervision. If there is little trust in tax officials, voters would prefer to retain more or less full control over tax policies, with changes requiring approval via referenda. In such cases, the tax

system adopted is likely to resemble the one outlined by Brennan and Buchanan (1978).

In practice, liberal democracies tend to have systems for adopting and revising tax systems that are in the intermediate category, which suggests that voters have intermediate levels of trust in their elected officials, and less trust in unelected panels of experts—especially those whose interests are not aligned with those of the typical voter (or democratic politics). That degree of trust is evidently commonplace within all liberal democracies.

Shared Interests in Transparency

A second area in which voters have shared practical interests is with respect to the transparency of tax systems. Voters share an interest in tax systems that help them minimizes mistakes when selecting among candidates and when selecting among private activities. Mistakes are costly—a matter that tends to be ignored in conventional optimal tax theory, which often proposes tax systems that are quite complex and often counter intuitive. Ramsay's theory implies that every good should have a different tax rate, which given all the tax shifting that takes place according to neoclassical economic theory implies that one's own tax burden would be nearly impossible to accurately estimate. No voter keeps a general equilibrium model in their head along with the requisite system-based econometrics and data sets. Mistakes at the voting booths from mistaken estimates and in their own private lives would be inevitable under such tax systems. Moreover, were such an "optimal" tax system ever adopted, it would tend to be undone through reforms that accord more with the economic intuitions of voters and elected officials.

The costly nature of mistakes when voting and making personal investment decisions implies that voters share practical interests in tax systems in which their tax burdens are easy to calculate. This requires several somewhat related properties that contribute to this feature of an ideal tax system for voters.

The first property that a tax system should exhibit is that tax systems should be easy to understand (e.g., lack complexity)—which eliminates most "optimal" tax conceptions from consideration. Tax systems should be such that a typical voter (not a genius) can accurately assess their own tax burdens. As a rule, these are tax systems in which the tax burdens fall mainly on the persons with direct payment obligations. The "checks" or "money transfers" made to

the tax authorities should have a high degree of correspondence with actual tax burdens. Examples include uniform sales taxes on widely traded goods, head taxes of various kinds, and perhaps taxes on land—e.g., taxes where the burdens tend to follow "check writing responsibilities.

A second feature that a readily understandable tax system should have is stability—it should not change year to year at random (or nearly so). One should not have to learn a new tax system every year. Unfortunately, tax systems tend to be unstable, because of majoritarian cycles. It is always possible to reallocate tax burdens in a manner that generates majority support for the reallocation. Thus, all tax systems tend to be unstable. However, it is more difficult to reassign cost shares under some tax systems than others. Administrative costs for changes in tax rates among groups of tax costs are higher for some systems than others, which makes them intrinsically more stable than others. In general, transparent taxes are taxes where it is difficult to discriminate among individual or classes of taxpayers. Examples include sales taxes, many user fees, and highway tolls— at least until recent innovations in technology.

Minimizing Agency Costs

A third feature that transparent (informationally efficient) tax systems should have is that agency costs should be easy for voters to detect (and punish). Both elected and unelected officials have the ability to undermine transparency by including special tax privileges for favored groups—either through the tax law itself or through decisions about how a tax system will be implemented.

Agency problems with respect to tax systems tend to be reduced by transparency, but also by frequent calculations of tax burdens. Transparency makes it easy to understand the tax code, frequent assessments of tax burden make it easy to understand whether one's understanding of the tax code is accurate or not. The best example is a sales tax—which one pays nearly every day on all sorts of goods. An income tax, in contrast, tends to be computed only once a year and only for one's own type of income. Taxpayers thus remain familiar with sales tax burdens associated with a fairly wide range of purchases, but income taxpayers tend to know only the taxes associated with their particular source(s) of income. The latter are thus more likely to include various loopholes that taxpayers will be unfamiliar with.

Other Shared Interests of Voters

In addition to the shared interests that voters have in informational efficiency, voters may also share a normative interest in particular types of taxation. Unfortunately, these may conflict with their interests in transparency. Normative interests often induce tradeoffs between practical interests in informationally efficient tax systems and other goals that voters might have regarding the fair or just distributions of tax burdens. For example, the norm that cost shares and tax burdens should reflect "ability to pay" tends to make taxes conditional on income and wealth, both of which are often difficult to measure, and thus efforts to apportion burdens in this way tends to undermine transparency.

When voters have internalized similar ideas about the nature of a "proper," "fair," or "just," tax system they may sacrifice some transparency and stability and accept some agency costs as the necessary price for advancing such normative goals. They may, for example, vote for progressive income taxes instead of sales taxes or in combination with them.

Unfortunately, there is no necessary reason why voter-taxpayers would share normative theories. Normatively optimal tax theories can be based on many normative theories—not just the utilitarian ones used by most optimal tax theorists. Examples of such families of optimal tax theories include: Lockean ones—grounded on natural rights and social contracts, utilitarian ones—with variations in weights and functional forms and hypotheses about economic consequences, and egalitarian ones—equal liberty vs equal outcome, or combinations of the two as in Rawls' theory of just taxation. Each of these normative approaches have implications for normatively optimal fiscal systems, and there may be few common grounds among them.

Shared Interests in Stability

When a super majority of voters share normative theories or reach normatively similar conclusion about taxes, the resultant tax system may not be as transparent as it might have been, but such agreement can solve or at least limit the cycling problems associated with discriminatory tax systems.

For example, suppose most voters agreed with Adam Smith' conclusions about a good tax system: "The subjects of every state ought to contribute toward the support of the government, as nearly as possible, in proportion to their respective abilities; that is, in proportion to

the revenue which they respectively enjoy under the protection of the state. The expense of government to the individuals of a great nation, is like the expense of management to the joint tenants of a great estate, who are all obliged to contribute in proportion to their respective interests in the estate." (*Wealth of Nations* p. 342)

Shared Normative Interests

If a majority of voters agreed with this description of an optimal tax system, it would clearly limit the domain of (normatively) acceptable taxes—here to proportional taxes, although there may still be some disagreements about the best tax base, the manner in which tax obligations should be shared would reduce or eliminate the political support for highly progressive or regressive tax system. This would reduce the likelihood of majoritarian cycles over tax systems and tax rates. And, insofar as such cycles still occur, it would tend to constrain the domain of cycling, thereby increasing the predictability of personal taxes through time.

The sharper the characterizations of the "ideal" agreed to, the less likely cycles would be to occur and the narrower would be the domain of taxes. Such ideals could, in principle, include some of the conclusions about tax systems reached by optimal tax theorists—although they do not necessarily do so.

A normative consensus would also tend to diminish agency costs, because government agents (both elected and unelected) would have internalized similar norms and those norms would inhibit (although not totally) their practical interests in tax systems that deviate from their own norms, norms that are shared with other officials and voters. Thus, a community with an ethos regarding taxes can shift away from transparent system of taxation without necessarily sacrificing a transparent tax system's low agency costs.

Unfortunately, when voters have normative interests, but their normative theories lead to different conclusions about "ideal" or "optimal" taxes, the stability of transparent systems tend to be undermined. Indeed, variation among norms make many of these problems more difficult to address in other ways, as with stabilizing laws and institutional refinements. Officials that disagree with the median voter's normative theory may feel morally obligated to try to replace that theory with his or her own, which tends to increase agency problems and undermine stability.

Such normative disharmony, thus, tend to bring informational efficiency back to the top of the shared interests of voters—but these are not likely to entirely dominate normative interests unless norms are weakly held, or their fiscal costs are obvious and very large.

VIII. Conclusion: Choosing among Tax Systems

Overall, this chapter has demonstrated that it is possible to think about "ideal" tax systems in a variety of ways. Individuals may choose them entirely with self interest in mind in which case they would tend to prefer systems that minimize their own tax burdens and that are generally transparent and stable. Alternatively, a voter's normative assessments of alternative tax systems may determine the system in place. Or, authority might be delegated to tax experts charged with finding the most normatively attractive system. Or, voters may retain some or all control over the selection of tax systems and make their own judgements about the normative and practical tradeoffs that inform their conclusions about "ideal" tax systems. Such systems will be less transparent than one based on practical consideration, but if their assessments of the impact of alternative tax systems are correct, the result will move their society in the direction that a majority of voters prefer to the alternatives.

Mainstream "optimal tax theory" may play a role in these assessments insofar as voters or politicians are influenced by their conclusions. However, a wide range of normative theories can be applied. When the resultant assessments of tax systems satisfy the Plott condition, a median voter would exist and an optimal tax theory that is relevant for democracies in which ordinary voters, rather than tax experts, determine optimal tax policies can be developed. Their views, of course, may be partly based at least partly the utilitarian ones used by most optimal tax theorists, although this is not necessarily the case.

The tax systems of liberal democracies suggest that tax systems are jointly determined by norms and pragmatic interests. They are less transparent than they would be if not influenced by norms regarding distributive justice, fairness, and property. And, evidently, there are points of agreement sufficient to limit cycles among "acceptable" tax systems to a relatively narrow domain—albeit one in which tax reforms take place nearly every year.

IX. Appendix: Additional notes on normative tax theories

A. James Buchanan (who won the 1986 Nobel Prize in economics while a professor at GMU, partly for his contributions to public finance, which tend to agree with Wicksell and Lindahl.

Buchanan argues that proper accounts of tax burden--should focus on net tax burden--that is, they should take account of the services financed by taxes as well as the taxes paid.

For example, if a person receives an especially valuable service from the government, it is possible that his or her "true" net tax burden is negative. Others who receive no services of value, might have positive net tax burdens.

Ideally, all citizens would bear "negative" tax burdens in the sense that each person should receive services that are considered to be more valuable than the taxes paid.

Buchanan points out that most Western governments are very productive in the sense that a good deal of the wealth produced in a given nation state is affected by property rights, civil law, and public services--as well the taxes used to finance them. These alone have benefits that exceed most person's tax costs.

B. Other normative principles of taxation come are rooted in shared cultural norms--often dealing with **fairness** (or equity).

i. Some argue that fairness requires all persons to pay be treated the same way under a tax system.

This notion of fairness tends to imply an equal cost sharing rule or a flat tax. Examples include a proportional tax on income which can be considered a flat "work hour" tax, where differences in marginal productivity are accounted for, as developed in Buchanan and Congleton (1998, Ch. 8).

ii. Others argue that taxation of persons should be based on their "ability to pay" or equal utility burdens from taxation.

This notion of fairness tends to imply progressive income taxes. For example, a "fair tax" might be one that caused all taxpayers should all sacrifice approximately the same "utility" (rather than net benefits) when they pay their taxes.

Since diminishing marginal utility implies that the marginal utility of money tends to be smaller for rich persons than poor persons, more money would be collected from rich persons than from poor persons. (This assumes that the MU of money curves are more or less the same. If some persons are better able to increase their utility using dollars than others, then they would bear lower burdens under this logic.)

This notion of fairness often plays an important role in policy debates over taxation in the US.

"Progressivity" if often argued to be desirable, while "regressivity" is often argued to be undesirable--although these ideas are not universally accepted among normative tax theorists.

iii. Perhaps the most extreme of the rational choice-based fairness analyses is that of John Rawls (1971, *A Theory of Justice*), who argues that fiscal packages should be designed so that the welfare of the least well-off person in society is maximized.

C. Some Useful definitions for characterizing and discussing tax schedules and also for discussion tax equity and/or the equity of fiscal systems from a utilitarian perspective:

i. A **progressive tax** is a tax whose average burden increases as the taxable base owned by an individual increases. [Such taxes often have marginal tax rates that increase with the base (increase with income), although not all progressive taxes have this property. Most income tax systems in industrialized countries are somewhat progressive.]

ii. A **proportional tax** is a tax whose average tax burden does not change with income. (Such taxes normally have a constant marginal tax rate, as true of most sales taxes and some income taxes. A flat (proportional) tax on income has the form: $T = tY$.)

iii. A **regressive tax** is a tax whose average tax burden falls with income. Such taxes often have declining marginal tax rates with ownership of the taxable base, however, not all regressive taxes have this property. An example of a regressive tax in the US is the social security tax--which has a cap on taxable income.

iv. [Instead of tax base, many analysts use income, which allows them to think in terms of "ability to pay" and apply utilitarian "fairness" norms to tax systems.]

D. A Few Useful Definitions and Relationships:

The **tax base**, B, is that which is taxed. (taxable income, sales of final goods and services, profits, property, gasoline, etc.)

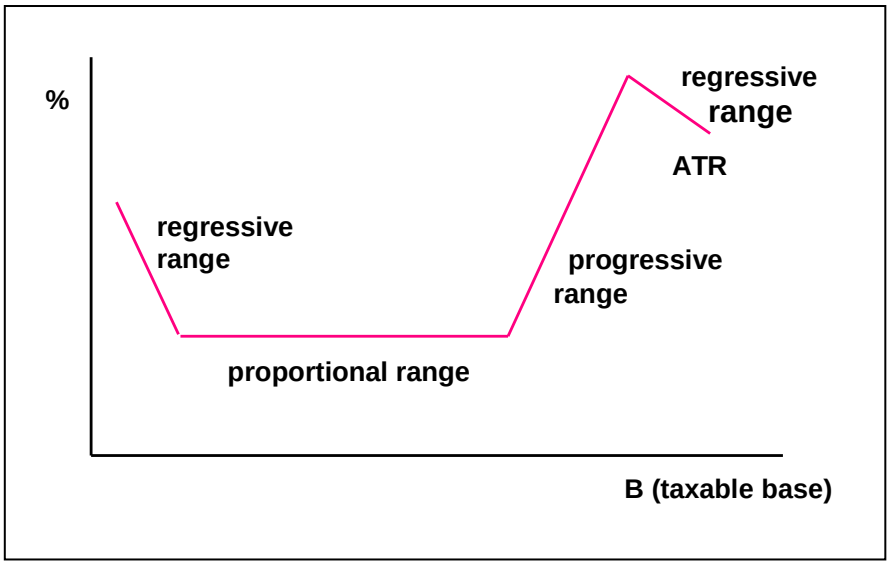
The **average tax rate** of a particular tax often varies with an individual's holding of the taxable base. If an individual pays tax Ti on a holding of Bi, his average tax rate is Ti/Bi. (If Ti = \$50 and Bi = 200, the average tax rate for this tax is 50/200 = 0.25 or 25%.)

The **marginal tax rate** of a particular tax is the change in taxes owed for a one unit increase in holdings of the taxable base, $\Delta T/\Delta B$. (So, if a taxpayer earning 50,000/year pays a tax of 10,000 and a taxpayer earning 50001 pays a tax of 10,000.50, his or marginal tax rate is 0.50/1 = 50%. Fifty percent of each additional dollar earned is taken from the "last" dollar of income earned by a taxpayer earning 50,000/year.)

In a **diagram of tax** schedules. If MTR is above ATR, then that ATR curve will be rising (the marginal tax rate will be pulling the average up). If MTR is below ATR, then the ATR curve will be falling (the marginal tax rate will be pulling the average down). If the MTR = ATR, the ATR will be neither rising nor falling.

Since individual decisions are determined by marginal cost and marginal benefits at various quantities, **it is the marginal tax rate rather than the average tax that affects taxpayer behavior**.

(Thus, one argument in favor of proportional, or indeed, regressive taxes, is that they may have smaller effects on economic activities than a revenue equivalent progressive tax.)



vi. (Peckman's estimates of the effective average and marginal tax rates faced by typical American taxpayers looks a bit like this odd tax schedule. This is mainly because Peckman considers reductions in transfer payments to the poor as their income increases as an implicit tax.)

(As an exercise try to determine what the marginal tax schedule that corresponds to this average tax schedule looks like.)

(Explain briefly why Peckman finds regressive ranges of taxation at both the highest and lowest ranges of income.)

E. The Importance of Marginal Tax Rates

The supply and demand diagrams of lecture 2 provide very useful ways to illustrate the burden of an excise tax, tariff, or other tax that can be represented in more or less "flat" per unit terms. However, they are less useful for examining the impact of more complex taxes such as a progressive income tax.

i. The taxes examined in the diagrams all had a constant tax rate, which implied that their average and marginal tax rates were essentially the same.

ii. In many cases, however, the marginal and average tax rates will differ, in which case it is the marginal tax rate rather than the average tax rate that is most important for predicting the impact of the tax on persons and markets.

iii. In such cases, one gains more insight into the effects of a tax by using indifference curve analysis or a bit of mathematics.

Consider the following decision setting in which Al can work to earn money for goods available only in markets or engage in leisure.

i. To simplify, assume that their H hours a day that can be worked and that Al is free to work as much or as little as he or she wants to.

ii. Also assume that work produces neither pleasure nor pain, but is s imply a means of obtaining the desired market basket of consumption goods, C.

iii. Leisure, L, is assumed to be a good subject to diminishing marginal returns as usual.

Hours worked are denoted W and the wage rate is w.

So, income is Y=wW.

The income tax schedule is T = t(wW) or T =t(w(H-L)) with $T_Y > 0$.

(Assume for convenience, that the entire burden of the tax is borne by Al.)

Given all this, we can write down the optimization problem that characterizes Al's labor-leisure choice:

$U = u(L, C)$

with $C = wW - T$ or $C = w(H-L) - t(w(H-L))$

Both leisure and consumption are ordinary goods subject to diminishing marginal utility and have positive or zero cross-partials.

$U_L > 0$, $U_C > 0$, $U_{LL} < 0$ and $U_{CC} < 0$ with $U_{CL} > 0$

To simplify the math a bit, note that one can substitute the budget constraint for C into the utility function.

This allows Al's utility can be written either entirely in terms of leisure (L)

or entirely in terms of hours worked W, if we also substitute for L = H - W.

In the latter case: U = u(H-W), wW - t(wW))

Differentiating with respect to W (the only control variable available to Al in this representation), we obtain the first order condition that characterizes Al's optimal workday:

W* satisfies: $U_L(-1) + U_C(w - T_Y w) = 0$

or $w(1 - T_Y) U_C = U_L + w(1 - T_Y)$

U_C is the marginal benefit of an hour worked (in utility terms) net of taxes, U_L is the marginal opportunity cost of working (also in utility terms). A utility maximizing person will work at the point where his marginal increase in income [$w(1 - T_Y)$] times the marginal utility of income equals the marginal utility of leisure.

Note that it is the marginal tax rate, T_Y, rather than the average or total tax rate, that affects Al's decision.

(In cases in which Al bears less than the full burden of the tax, it will be his or her effective marginal tax rate that affects behavior.)