Casey B. Mulligan

Notes on the Behavioral Implications of Tax Distortions

Here we set up a simple model illustrating some of the prototypical results in public finance regarding the effect of taxes on behavior. These results relate to the economic equivalence of various kinds of taxes, the wealth and substitution effects of taxes, and deadweight costs.

Throughout these notes, we consider an economy with two mutually exclusive composite "goods," which we call "consumption" c and "leisure" l. These terms should not usually be interpreted literally, because the important distinction for our analysis is that "consumption" is associated with taxable activities, "leisure" with nontaxable activities, and that these two goods are not perfect substitutes. For example, leisure $\in [0,1]$ does not necessarily refer to the fraction of the week that a person is not on his job but, depending on the application, could also refer to the fraction of time the person works in the nontaxable sector, or the fraction of household members who have a paid job, or even the fraction of a person's lifetime in which he is employed. As discussed in an earlier lecture, we literally assume that taxpayers do not obtain any utility from the taxes they pay, which is meant to capture the more general situation where a person prefers the marginal taxdollar in his pocket rather than at the Treasury.

In the prototypical model, employers pay employees a "wage rate" w per unit time. The bulk of our analysis pertains to a wage income tax at rate τ paid by employees, so an employee's budget constraint is:

$$c = w(1-\tau)(1-l) + \nu + p \tag{1}$$

because taxes are paid in the amount $\tau w(1\text{-}l)$. $\nu \geq 0$ is a lump sum transfer (namely, paid by the government to the employee regardless of his behavior). p is "property income" such as dividends. Including it allows us to consider questions of intertemporal behavior by taxpayers, and intertemporal features of tax rules.

I. Equivalence Results

Most of these notes focus on a tax that can be literally interpreted as a labor income tax. This is relevant in part because so much government revenue around the world comes from payroll taxes and from personal income taxes (most personal income is labor income). A lot of revenue also comes from "value-added" and "sales" taxes, but much of the labor income tax analysis also applies to these consumption taxes. Furthermore, even if we ignore the distortionary effects of raising government revenue, the process of spending revenue can also have many of the same distortions as a labor income tax. Section I of our notes uses the assumptions above to derive these and other relations between consumption, labor income, and savings taxes, and means-tested transfers.

I.A. *Consumption vs. Wage Taxes*

Consider, for example, replacing the wage income tax above with a consumption expenditure tax levied at rate t per dollar of consumption expenditure. The budget constraint becomes:

$$(1+t)\, c = w\,(1-l) + \nu + p$$

dividing through by $(1+t)$, we have:

$$c = \frac{w}{1+t}(1-l) + \frac{\nu + p}{1+t} \tag{2}$$

Comparing (1) and (2), we see that both consumption and wage income taxes have the effect of reducing the slope of the employee's budget constraint in the $[l,c]$ plane. Furthermore, a wage income tax rate τ is equivalent to a consumption tax at rate t if $\nu+p = 0$ and $(1-\tau)(1+t) = 1$. An employee with $\nu+p > 0$ prefers the wage tax, while an employee with $\nu+p < 0$ prefers the consumption tax, which is one reason why it is said that the elderly prefer payroll taxes rather than sales taxes.

I.B. Subsidies vs Taxes

Suppose that the government guarantees a minimum income in the amount ν to all persons, and does so by levying a tax at rate τ on all wage income earned beyond the amount ν. In other words, a person's tax liability is $\tau w(1-l)-\nu$, which turns out to be a subsidy for all persons with $w(1-l) < \nu/\tau$. The budget constraint is therefore:

$$c + \tau w(1-l) = w(1-l) + \nu + p$$

which is the same as (1). In words, both a wage income tax and an income guarantee have the effect of reducing the slope of the employee's budget constraint in the $[l,c]$ plane.

I.C. Nominal Tax Liabilities

Does it matter whether the employer or the employee has to pay the tax? One additional assumption is required to deliver a simple answer. Namely, let f' denote the cost of leisure time to employers (again, depending on the application, this can refer to the number of persons working at the firm, the hours worked by each employee, etc.) in terms of foregone revenue and τ_f denote the rate at which employers pay taxes on the wage income they pay to their employees. In a competitive labor market with profit maximizing firms, the wage rate paid to employees will satisfy:

$$w\,(1+\tau_f) = f'.$$

The LHS denotes the two expenditures incurred by an employer when an employee reduces leisure – funds go to the employee at rate w and to the Treasury at rate $\tau_f w$. If this were greater (less) than f' the employer would ask his employees to take more (less) leisure.

The slope of the employee's budget constraint is $w(1-\tau_e)$ which, by the condition above, is $(1-\tau_e)f'/(1+\tau_f)$. If we define the "marginal tax rate" (MTR) to be the percentage gap between the employee's budget constraint and the marginal rate of transformation f', we have:

$$MTR = \frac{\tau_e + \tau_f}{1 + \tau_f} \tag{3}$$

In other words, a *x*% wage income tax paid by employers is not the same as a *x*% wage income tax paid by employees because, given the marginal rate of transformation, the tax levied on the employer lowers the wage rate while the tax levied on employee does not (see also Barro and Sahasakul 1986, p. 557).

Consider some international examples of payroll taxation. In 1995, Singapore had a payroll tax for which employees owed 20% of wage income and employers owed another 20%. In other words, $\tau_e = \tau_f = 0.20$. From our formula (3), this implies a marginal tax rate of 0.333. In other words, under our assumptions, Singapore's payroll tax would be no different if the employer portion were repealed and the employee rate increased to 33.3%.

In 1975, Chile had a payroll tax for which employees owed 7% of wage income and employers owed another 49%. To the layman, Chile's system may seem more distortionary than Singapore's because 49% exceeds 20% + 20% – and this doesn't even mention Chile's 7% charged to employees. However, our formula (3) shows how the 49% tax on employers is by itself equivalent to a 32.9% tax on employees and hence very much like Singapore's system. When we add on the 7% charged to employees, Chile's 1975 *MTR* is 37.5% which is only slightly higher than Singapore's. We return to this point later when we discuss Social Security systems around the world, because heavy tax rates on employers are prevalent and because the division of the nominal tax burden between employer and employee varies a lot from country to country.

There is at least one interesting case where these formulas may not apply – namely, when there is a binding minimum wage. To see this, notice how our derivation required $w = f'/(1+\tau_f)$ so that a high tax rate for employers decreases w, increases f', or some combination keeping profits constant. With a binding minimum wage, higher τ_f can only increase f' or decrease profits. To the extent it is the latter, there may be a gap between w and $f'/(1+\tau_f)$, and a higher employer tax rate serves to close that gap and thereby drive a lesser wedge between employee's budget constraint and the marginal rate of transformation f'.[1]

II. Wealth and Substitution Effects of Taxes – "Aggregate" Analysis

Taxes are expected to have offsetting wealth and substitution effects. Taxes reduce taxpayers' purchasing power and are therefore expected to cause them to spend less on all normal goods, including "leisure" or the "untaxed activity." Taxes also make the untaxed activity relatively cheap, a substitution effect that by itself encourages the untaxed activity. Here we formally derive these results, discuss how they are (somewhat differently) applied

[1] See also the material from Steve Trejo's dissertation, such as Trejo (1991).

in aggregate and microeconomic policy studies, and amend the analysis to allow for a "double substitution effect." Throughout we keep in mind how the wealth-substitution effect decomposition is important for answering two important public finance questions:

- what are the behavioral effects of public policy?
- what are the efficiency effects of public policy?

II.A. Efficient Allocations

To begin, we borrow the notation from above, ignore net savings, and interpret the taxpayer as the economy's "representative agent." In addition to the utility they obtain from consumption and leisure, taxpayers also enjoy (or dislike – we put no restrictions on the derivatives of this part of the utility function) government consumption activities, which occur in the amount g. In summary, there are many identical taxpayers, each with utility function:

$$u(c, 1-n) + J(g) \tag{4}$$

where u is strictly increasing in both arguments and jointly concave, and we assume that consumption and leisure are normal goods in the relevant range (ie, the cross-derivative of u is not too negative in the relevant range). We use n as a shorthand for one minus leisure, and can interpret it as the amount of time devoted to production of taxable goods.

g, as well as transfers ν (to be introduced later), are exogenous variables, taken as given by "the planner" and by individual agents.[2] J is exogenous, and additively separable from u, so we can (and will) ignore it when characterizing efficient and equilibrium behavior.

The economy's resource constraint is simply:

$$f(n) = c + g \tag{5}$$

where output f in the taxable sector is a strictly increasing and concave function of labor supplied, and $nf'(n)$ also increases with n. There also may be lump sum taxes and transfers in the economy, where the government steals an amount from some (all) agents to finance spending to some other (all) agents, and the amounts taxed from, or transferred to, an agent are independent of his behavior. But lump sum taxes and transfers do not affect the resource constraint (5).

Efficient allocations can be described by a planner's problem:

---

[2]Analyzing the optimal policy is interesting, but we defer that until later in the course.

$$n = \underset{x}{\operatorname{argmax}}\, u(f(x) - g, 1 - x)$$

$$c = f(n) - g$$

We can derive the effect of government consumption on private behavior by graphing the planner problem in the [n,c] plane.

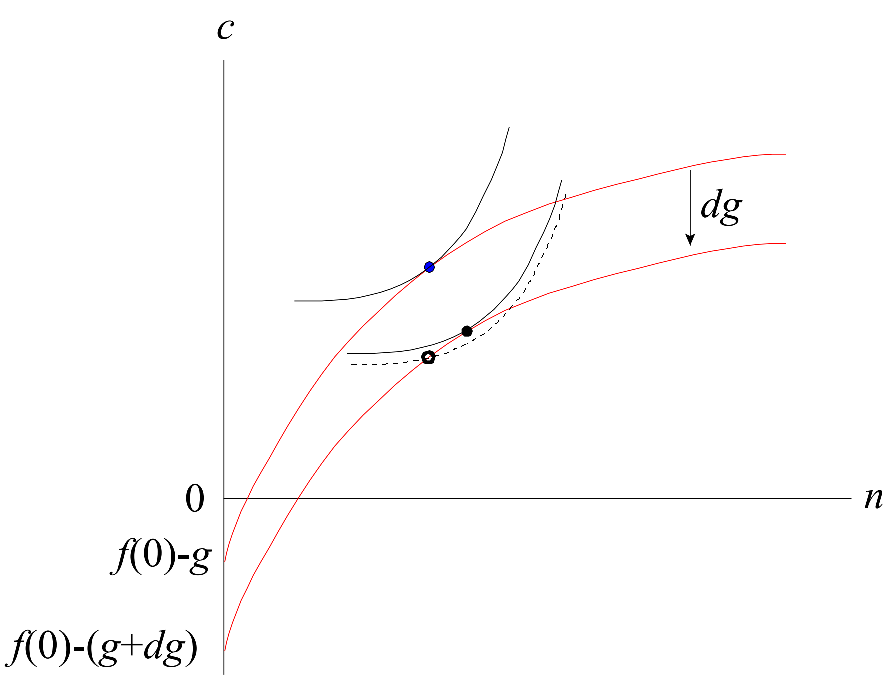

Figure 1 Policy Effects in an Efficient Economy

The red concave curves are $f(n) - g$, and shift down in a parallel fashion in response to more government consumption as in Figure 1. A black convex curve is an indifference curve, and is tangent to the red curve at the optimal allocation, with the original allocation marked as a blue circle and the new one as a black one. The normality of leisure and the convexity of preferences imply that, if work time were unchanged, the marginal rate of substitution would decline and would no longer be equal to the marginal product of labor, as shown by the dashed indifference curve and the hollow circle. So time worked must increase and, by a similar argument, consumption must fall.

II.B. Equilibrium Interpretations of Efficient Allocations

In essence the policy variable g is just a parameter describing production sets and preferences, which are convex, so we can apply the second welfare theorem to prove that the efficient allocations can be implemented as competitive equilibrium allocations. It is important to see the mechanics of this. First, define a competitive equilibrium:

Definition Given a policy $\{g, \nu\}$, a competitive equilibrium with lump sum taxes is a list of scalars $\{p,c,n,w,T\}$ such that:

(i) given $\{p,w,T,\nu\}$, $\{c,n\}$ maximize (4) subject to:

$$wn + \nu + p - c - T = 0$$

(ii) The resource constraint binds:

$$f(n) = c + g$$

(iii) given w, p and n solve:

$$p = \max_n f(n) - wn$$

(iv) $\{g,\nu,T\}$ balances the government budget constraint:

$$g + \nu = T$$

In other words, atomistic households are maximizing their utility and firms are maximizing profits given prices and policy. In particular, each household owns a share of a firm, whose market value is p, and faces a wage rate, tax, and transfer. They buy or sell shares of firms in order to achieve the desired consumption and work time. Managers of the firm demand factors that maximize cash flow. Although each household considers buying or selling shares of firms, an equilibrium is not consistent with any net sales in the aggregate. In other words, the value of firms adjusts so that the average household is willing to hold exactly his share.

The first order conditions for households and firms are:

$$\frac{u_l(c, 1-n)}{u_c(c, 1-n)} = w = f'(n)$$

We see above, and know from the second welfare theorem, that any efficient allocation can be supported as a competitive equilibrium. In particular the equilibrium marginal rate of substitution is equated to the marginal product of labor. If we are given an efficient allocation $\{c,n\}$, we can compute an equilibrium wage by plugging efficient allocations into the first order conditions above. We have already evaluated the effects of policy on efficient allocations, so now we can evaluate the effects of policy on equilibrium prices.

There is a stable, downward sloping, derived demand for labor, implicitly defined by the firm's first order condition (6):

$$w = f'(n) \tag{6}$$

More government consumption raises work, so it follows from (6) that the pretax wage rate falls. In other words, the primary effect (in the $[n,c]$ plane) of government spending and taxes is an adverse wealth effect, and the extra work that derives from this decreases wages. The reduced wage by itself might reduce work (through a substitution effect on labor supply), but this cannot fully offset the wealth effect because it is a reaction to the behavior deriving from the wealth effect.

II.C. The Government Spending "Multiplier"

Politicians, journalists, and some economists have explained why they think government spending has a multiplier effect. If the government, they argue, purchases $1 million worth of goods from firm ABC, the owners and workers at that firm have $1 million more available to spend on their own wants and needs. For example, they might decide to spend an additional $0.5 million purchasing goods from other firms, who we'll call DEF. The firms and workers at firm DEF now have $0.5 million more available to spend, of which they may decide to spend $0.25 million from other firms, who we'll call GHI. The owners and workers at firm GHI spend $0.125 million, and the process continues.... The aggregate effect of the initial $1 million government purchase is to increase public and private spending by a total of $2 million:

$$1 + 0.5 + 0.25 + 0.125 + \ldots = 2$$

In other words, government spending is said to *increase* private spending. The total (public + private) effect on spending of the initial $1 million is greater than $1 million, so government spending is said to have a "multiplier effect". Because the magnitude dy/dg of the effect is called "government spending multiplier," the politicians, journalists, etc., are saying that the government spending multiplier is greater than one.

The government spending multiplier is less than one in our model. Equivalently, government spending in our model reduces private spending.[3] This can be proved in two ways. On the demand side, consumers feel poorer due to the taxes they pay, and compensate for this lost income in part by reducing their demand for (normal) private consumption goods. On the supply side, both private and public goods must be produced with labor, and consumers are not willing to supply so much extra labor as to provide additional demand for from the private and public sectors. Thus private spending must fall.

II.D. Equilibrium Allocations with Labor Income Taxes

Now we show how the economy's response to a change in government consumption, or transfers, in theory ought to be very different when the spending is financed with a labor income tax rather than a lump sum tax. The first step in the analysis is to describe how the labor income tax works. In our model, the government simply levies a tax on each household equal to a fraction τ of its labor income for that period, wn. If a household chooses

[3] Remember that GDP is the sum of private and public spending.

to work less (or more), it will have less (or more) labor income and owe less (or more) in taxes. Second, we define a competitive equilibrium with labor income taxes.

<u>Definition</u> Given a policy $\{g, \nu\}$, a competitive equilibrium with labor income taxes is a list of scalars $\{p, c, n, w, \tau\}$ such that:

(i) given $\{(1-\tau)w, \nu+p\}$, $\{c,n\}$ maximize (4) subject to:

$$(1-\tau)wn + \nu + p - c = 0$$

(ii) The resource constraint binds:

$$f(n) = c + g$$

(iii) given w, p and n solve:

$$p = \max_n f(n) - wn$$

(iv) $\{g, \nu, \tau, w, n\}$ balances the government budget constraint at each date:

$$g + \nu = \tau wn$$

As with lump sum taxes, atomistic households are maximizing their utility and firms are maximizing profits given price and policy sequences. But notice how (selling) households face a different labor price $(1-\tau)w$ than do the (buying) firms, w. In particular, an additional unit of labor yields revenue $(1-\tau)w$ to the household while costing the firm w – the government gets the difference.

II.E. The "Laffer Curve," and the Existence and Uniqueness of an Equilibrium

Notice from (iv) that the rate of labor income taxation will never be negative, because at best the government has no spending. Also notice three properties of household behavior: (a) no tax revenue is collected if $\tau = 0$, (b) some revenue may be collected if $\tau \in (0,1)$, and (c) no tax revenue is collected if $\tau = 1$. (a) and (b) follow trivially from tax revenue's being the product τwn. (c) follows because a household has no incentive to work when $\tau = 1$, so that $n = 0$, which means that $\tau wn = 0$. It should not be surprising that there is some revenue maximizing tax rate between zero and one and an associated maximum revenue and, if government spending $g + \nu$ exceeds this maximum, no competitive equilibrium allocation exists even if there exist feasible allocations (as defined in Section I).[4] We suppose that $g + \nu$ is small enough that a competitive equilibrium exists.

It should not be surprising that there may be multiple tax rates yielding the same revenue. For example, there are at least two tax rates yielding zero revenue, namely 0 and 1. We shall consider only the equilibrium with smaller tax rates. Heuristically, this means we

[4]The relationship between tax revenue and the tax rate is sometimes called the "Laffer curve."

consider equilibria on the upward sloping part of the Laffer curve.

II.F. The Tax Wedge and the Effects of Transfers

To characterize equilibrium allocations we write down the first order conditions for households and firms:

$$\frac{u_l(c,1-n)}{u_c(c,1-n)} = (1-\tau)w < w = f'(n) \tag{7}$$

Here we have an important difference from the lump sum tax case: the marginal rate of substitution is less than the marginal product of labor. The gap between marginal product and marginal rate of substitution is sometimes called the "tax wedge."

The tax wedge is one of the most basic concepts in public finance because all distortionary taxes drive a wedge between some marginal rate of substitution and its corresponding marginal rate of transformation. The labor income tax wedge is seen in our model where the competitive equilibrium $\{c,n,w,\tau\}$ can be characterized according to four equations: the two intratemporal first order conditions (ie, equations (7)), the resource constraint (5), and the government budget constraint. Graphically, these four equilibrium conditions can be seen in a single diagram, as in the left Figure 2.

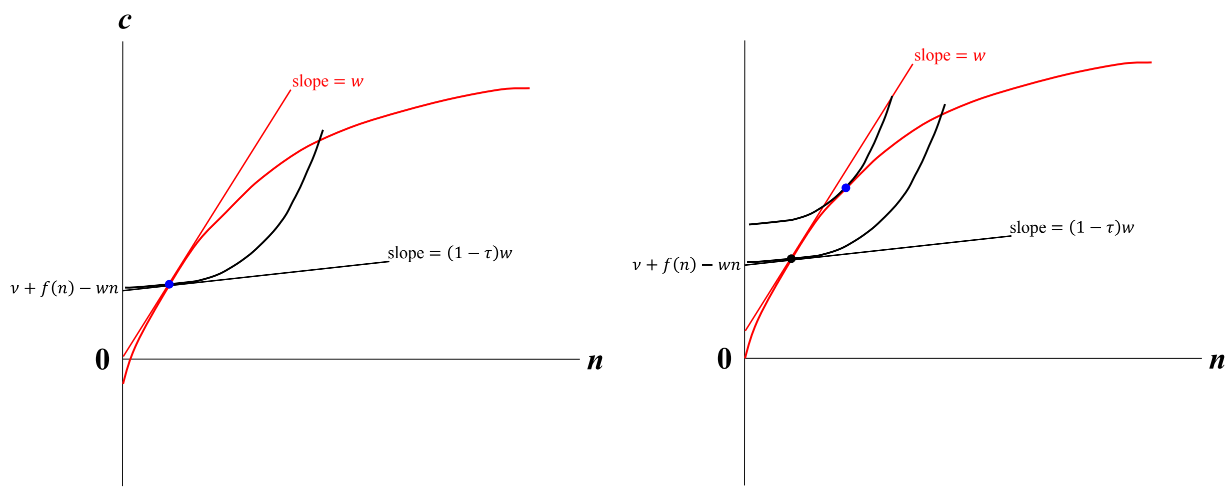

Figure 2 Graphical Analysis of the Tax Wedge
(left: equilibrium with positive g,ν,τ; right: the effect of transfers)

Equilibrium consumption and work time are shown as a solid dot, and are feasible because the dot lies on the red schedule $c = f(n) - g$. An isoprofit curve for a firm is shown as a red line, and has slope equal to the pretax wage, w. The equilibrium is consistent with profit maximization because the isoprofit line is tangent to the production set $f(n)$. A household budget constraint is shown as a solid black line; its slope is the after-tax wage $(1-\tau)w$ and its height is the sum of lump sum transfers ν plus profits $f(n)-wn$ received (and taken as given) by the household from firms. Optimization by households is shown as tangency between the black indifference curve and the household budget constraint. Government budget balance can be proven by using the height and slope of the household budget constraint to compute consumption: $c = [\nu + f(n) - wn] + n(1-\tau)w = \nu - \tau nw + f(n)$. Since the solid dot is also on the red schedule, we also have $c = f(n) - g$, which implies $\nu + g = \tau nw$.

We immediately see from the left Figure 2 how the competitive equilibrium is inefficient. Furthermore, the inefficiency derives from the wedge between prices faced by households and firms, and that the efficient allocation involves less leisure time. Lump sum transfers do not affect equilibrium allocations when they are financed with lump sum taxes, but do when financed with distortionary taxes.

Now consider an initial policy $\{g, \nu\}$, and increase only ν. Assuming that the government is not already at the top of its Laffer curve, such an increase is feasible, and requires a higher τ. This policy change increases the wedge between prices faced by households and firms, and must decrease consumption and work time. This is most easily seen when we set g and the initial transfer ν equal to zero, as in the right Figure 2. In this case, the initial equilibrium is efficient, and shown as a solid blue dot. The final equilibrium is shown as a solid black dot. Since its tax rate is positive, the final equilibrium must have the property that the indifference curve is flatter than the production possibility frontier – this is the tax wedge – which can only occur with less consumption because of the convexity of preferences. The allocation is less efficient, has less work time, has a higher pretax wage, but a lower after-tax wage.

It is important to understand the "Logic of Collective Action"[5] involved in Figure 2, because it is the cornerstone of modern public finance theory. There are many workers in the model economy and, even though they are identical, they make their decisions independently. In equilibrium, all of those decisions are the same, but each individual worker decides how much to work (a) assuming that the taxes he pays will mainly serve to increase transfers for other people and (b) not valuing transfers to others nearly as much as he values income for himself. Hence, the ideal for any individual is for him to equate his date t intratemporal marginal rate of substitution (MRS) to $(1-\tau)w$ while all other workers equate their MRS to w. In terms of the figure below, this means he would consume at the hollow dot – because the others would be working hard and generating a lot of government revenue – while all others consumed at the blue dot! But all workers (correctly) reason that way, and the result is less transfers, or a higher tax rate, for all.

[5]The title, and careful exposition of the logic, is due to Olson (1965).

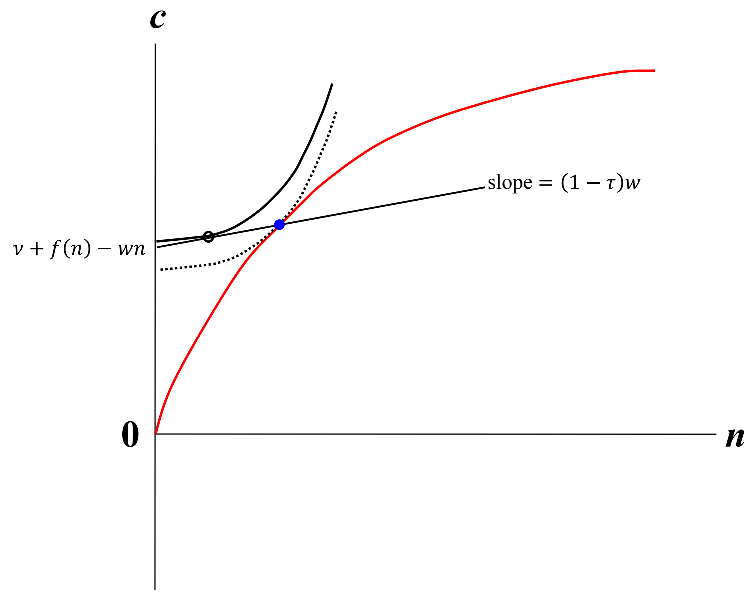

Figure 3 The Logic of Collective Action

IV.G. The Laffer Curve

As we vary v, we vary the equilibrium work time n and tax rate τ. We can parametrically derive a Laffer curve by (a) varying n, (b) using preferences and technology to compute the tax wedge, and thereby τ, and (c) computing the transfer amount v that would induce n as a competitive equilibrium amount of work time. Figures 4 do so.

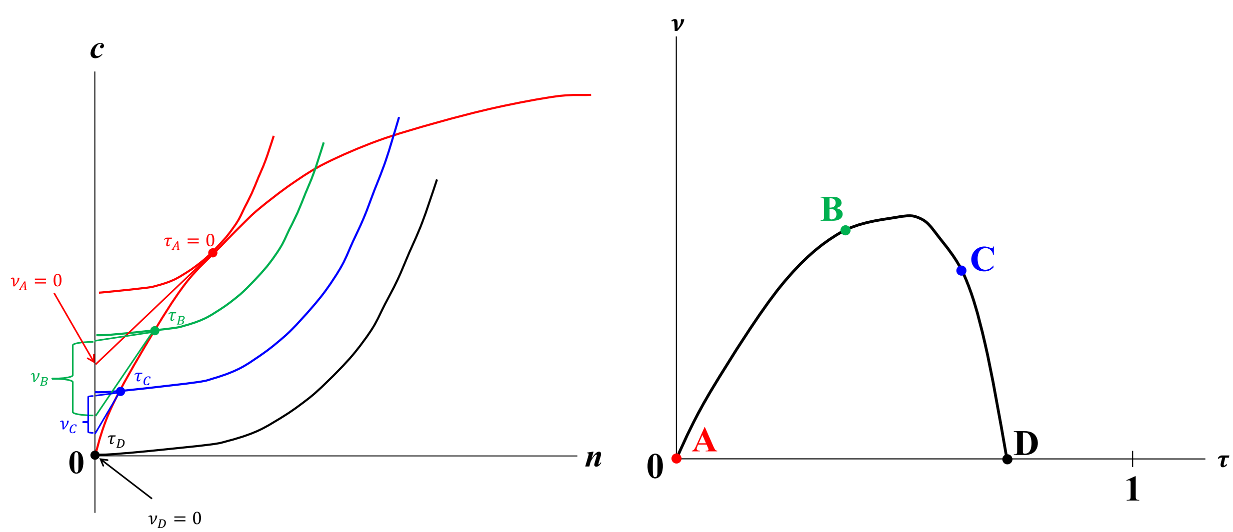

Figure 4 A Laffer Curve

On the left, we progressively reduce n by moving from point A, to B, to C, and finally to D

(marked in red, green, blue, and black, respectively). The left diagram permits computation of the tax rate at each point – it is the percentage difference between marginal product and marginal rate of substitution (MRS). ν can also be computed at each point by computing the intercepts of lines with slope equal to f' and MRS at each point, and taking the difference.[6] We then plot τ vs. ν in the right Figure, and have a graph of tax rate vs. tax revenue – a Laffer curve. Notice from the left Figure how revenue is zero for allocation A because f' and MRS are identical, and zero at allocation D because $n = 0$. We clearly see from the left Figure how $\tau_A < \tau_B < \tau_C < \tau_D$, while $0 = \nu_A < \nu_B > \nu_C > \nu_D = 0$. Filling in points in between points A, B, C, and D would trace out the full Laffer curve, which might be like that is drawn in black in the right Figure.

We see above how competitive equilibria are typically not unique. Comparing Figures 4 and the right Figure 2 reveals how not appreciating the nonuniqueness can lead to different, and probably uninteresting, comparative statics. Suppose we want to know the effect of increasing ν in an economy with $\nu = g = 0$. If we begin with the "wrong" equilibrium, namely the allocation D in Figures 4, we would conclude that *more* transfers increase work via *decreasing* the tax wedge!

II.G. Labor Income Tax-Financed Government Spending

Consider for a moment the effect of financing additional government consumption by reducing transfers by an equal amount. Graphically, this might be most easily studied by beginning with the black allocation in the right Figure 2, where we have positive transfers and zero government consumption. Greater government consumption shifts the production frontier downward in a parallel fashion by the amount dg, as in Figure 5. If the tax rate and work time were unchanged, the household budget constraint would shift downward by the same amount, to the black dashed line. However, we know from the normality of leisure and the convexity of preferences that the marginal rate of substitution is lower at the hollow circle than at the solid black circle, so that hollow circle cannot be a competitive equilibrium. Work time must increase, and consumption and the pretax wage fall. With more work time, the tax rate must be lower to balance the budget, although the total effect on the after-tax wage is negative. Hence financing government consumption by reducing transfers increases work via a wealth effect as it did in the lump sum economy, but also reduces the tax wedge and both pre- and after-tax wages, as seen by comparing the solid black and green circles in Figure 5.

---

[6]The intercept of the line with slope f' is $f(n)$-wn, which is firm profits. Remember that the intercept of the household budget constraint is the sum of transfer and profits, so the difference between intercepts is the transfer.

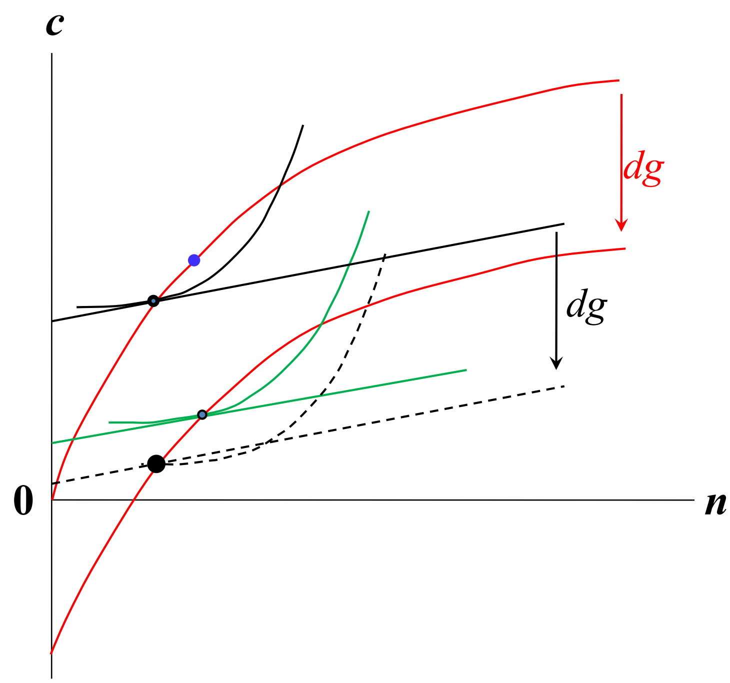

Figure 5 Government Consumption with Distortionary Taxes

To compute the effect of financing additional government consumption with a labor income tax, we need only to add the effect of financing additional government consumption with transfer reduction (the difference between the solid green and black dots in Figure 5) with the effect of increasing transfers financed with a labor income tax (the difference between the solid blue and black dots in Figure 5 or the right Figure 2). Clearly we have less consumption and a larger tax wedge, but we cannot sign the effect on work time: we have the adverse wealth effect of the government consumption going against the substitution effect of the tax wedge. Nor can we sign the total effect on the pretax wage, although we have a stable labor demand curve (6) so w and n must move in opposite directions. The magnitude of each effect depends on leisure demand's income elasticity (ie, how rapidly marginal rates of substitution increase with c) and wage elasticity (ie, the degree of convexity of preferences). The after-tax wage falls.

II.H. The Double Substitution Effect

Regarding transfers and the quantity of work, one should recognize that transfers are often not lump sum. For example, Social Security and other old age benefits are paid as a declining function of the elderly beneficiary's earnings. A benefit b would be paid to those elderly with zero earnings, and then benefits would be reduced with earnings at rate β (sometimes known as the "benefit reduction rate"). In other words, the transfer is:

$$v = b - \beta wn \tag{8}$$

Plugging (8) into the household budget constraint and rearranging terms, we have:

$$(1 - \tau - \beta)wn + b + p - c = 0$$

which is analytically identical to the model we analyzed in the main text, but we have two different interpretations:

- the marginal tax rate is the sum of labor income tax and benefit reduction rates, $\tau + \beta$
- the lump sum transfer is not the amount appearing in the government budget, ν, but the amount b beneficiaries would receive if none of them worked

An important book computed the sum $\tau + \beta$ for Social Security beneficiaries in eleven countries (Gruber and Wise 1999), which I graph versus the elderly labor force participation rate in Chart 1. We see how these rates are much higher than the payroll tax rates τ and are highly correlated with elderly labor force participation.[7] The U.S. House Ways and Means Committee (1996, Table 8-3) explains how to combination of labor income tax and benefit reduction rates yield very weak incentives to work for (non-elderly) mothers with low earnings potential. I summarize some of their calculations below in Table 1.

[7]Another difference between Social Security benefits and transfers in our model, is that Social Security benefits are paid only in old age, apply only to the old the large marginal tax rates shown in Chart 1, and hence have a life cycle substitution effect that we do not see in the right Figure 2. Hence, we expect Social Security benefits to decrease elderly work substantially more than they decrease average work.

Table 1: Earnings and "Disposable Income"
for a Mother with Two Children and Day Care Expenses

annual earnings, $	annual disposable income, $	"explicit" tax rate, %	"explicit+implicit tax rate," %
2,000	9,349		
4,000	9,916	7.65	71.65
5,000	10,199	7.65	71.70
6,000	10,483	7.65	71.60
7,000	10,766	7.65	71.70
8,000	10,986	7.65	78.00
9,000	11,785	7.65	20.10

Source: U.S. House Committee (1996, Table 8-3). Refers to January 1996, for a PA mother who has been on her job for at least 4 months. "Disposable Income" is earnings plus EITC, AFDC cash, the face value of food stamps, and less payroll and personal income taxes, and work expenses. "Explicit Tax Rate" is the payroll+income tax change divided by the earnings change. "Explicit + Implicit Tax rate" is one minus the disposable income change over the earnings change

The first column is earnings, which we might interpret our model's nw. The second column is the sum of earnings and various transfers net of taxes, which we may interpret as $(1-\tau)wn+\nu$. The third column is the rate at which payroll and personal income taxes accumulate with earnings.[8] The final column is calculated from the first two, reflects how the sum $(1-\tau)wn+\nu$ increases at rate less than one with wn, and can be interpreted as $\tau+\beta$. We can disagree with some of the details of the House Committee's calculations, but the Table illustrates how transfers can act like taxes, and are quantitatively significant.

Finally, we see from equation (8) that government benefit formulas have automatically built in a negative relationship between transfers and beneficiary work, even if labor supply were completely inelastic to taxes. Holding constant b and β, ν must increase when there is more retirement because more beneficiaries are eligible for the full benefit b.

II.I. Consumption as the Scale Variable

The marginal rate of substitution function provides another way to decompose the behavioral effects of wages and taxes. Namely, we can use the taxpayer first order condition $w(1-\tau)=M(c,1-n)$ to calculate labor supply as an implicit function of consumption and the after-tax wage. Figure 6 graphs this relation in the $[n,w]$ plane for a given level of

[8] In practice this is only the payroll tax, because the House Committee calculates that a mother of two earning less than $10,000 would not be liable for a personal federal or state income tax.

consumption.

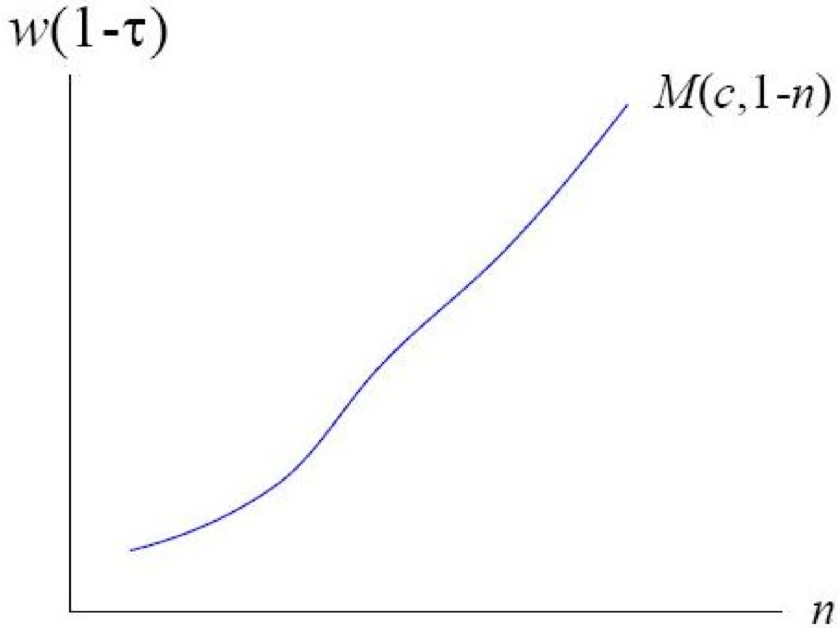

Figure 6 Labor Supply and the Marginal Rate of Substitution

This curve slopes up under the assumption that consumption is a normal good. It shifts up with consumption under the assumption that leisure is a normal good. Tax rates affect the after-tax wage (both directly and through the equilibrium effect on pre-tax wages) and consumption, and thereby labor supply. We might decompose the total effect into a shift of the M curve in the $[n,w]$ plane and movements along that curve, with the former a "wealth effect" and the latter a "substitution effect." We extend this approach in later lectures, and apply it to some questions in macroeconomics and economic history.

II.J. Behavioral and Efficiency Effects of Taxes

Now we can return to two of the key public finance questions: what are the behavioral and efficiency effects of taxes? When the taxes are lump sum, we see in Figure 1 that they affect behavior but not efficiency on the consumption-leisure margin. When the tax is levied on labor income, we could have a large effect on behavior and a small effect on efficiency, or a small effect on behavior and a large effect on efficiency. To see an example of the former case, consider Figure 5 and imagine that the indifference curves are nearly Leontief. Here taxation could significantly change labor supply, but the extra utility that could be gained by switching to a lump sum tax would be nil. In other words, the wealth effect is large relative to the substitution effect, so that the total behavioral effect is large, and the efficiency effect small. Figure 5 also provides an example of the latter case, if we take it quite literally in the sense that the solid blue dot is exactly above the solid green dot. Here to the total effect on labor supply is nil, but there is clearly an effect on efficiency because the indifference curve crosses the production set at the solid green dot. In other words, there is a significant substitution effect, and it is canceled by the wealth effect in terms of labor supply. The result that behavior is not isomorphic to efficiency is why decomposing tax-induced behavioral responses is an important topic in public finance.

III. Wealth and Substitution Effects of Taxes – The Consumer Theory Perspective

The consumer theory perspective treats the wage and tax rates as parameters and examines their effect on taxpayer behavior, without much attention to which changes in those parameters are consistent with budget balance or market clearing. Doing so may be useful, for example, when comparing differently-taxed groups in the same labor market (so the pretax wage is a common parameter), or when analyzing the responses to changing tax rules for only a small group of persons who have a negligible effects on the market and the government budget. Even when government budget balance and market clearing need attention in the analysis, the consumer theory approach can be useful as an intellectual exercise for understanding a piece of the logic of equilibrium models.

Analytically, the consumer theory approach to labor supply solves the following program:[9]

$$\max_n u(p + (1-\tau)wn + \nu, 1-n)$$

Notice how the consumer theory version is an algebraic special case of the "aggregate" analysis above – namely, the special case where government consumption g is adjusted to balance the government budget in response to any change in the parameters of the consumer problem (τ, ν, w, or p). It follows from our aggregate analysis that a larger tax rate has an ambiguous effect on consumer behavior because there are wealth and substitution effects in opposite directions.

We can learn more about the total effect of taxes on labor supply by using the Slutsky equation:

---

[9]In this analysis, nonlabor income ($p+\nu$) is treated as a separate parameter from the after-tax wage $(1-\tau)w$. This treatment maybe of limited usefulness in a lot of applications because nonlabor income itself depends on the after-tax wage. Consider, for example, the analysis of married women. A married woman's decision to work could be modeled as an individual decision like that above, where the key parameters are her after-tax wage and her "nonlabor income," important components of which are property income and the "subsidy" she receives from (or pays to) her husband and other family members. But it seems unlikely that we would observe a situation where, say, her after-tax wage and labor supply changed, but these components of her nonlabor income were unchanged. Perhaps the more useful analysis would aggregate across persons or over the life cycle, but we leave this concern to the side for now and return to it in a later lecture.

$$\frac{\tilde{w}}{n}\frac{dn}{d\tilde{w}}\bigg|_p = \frac{\tilde{w}}{n}\frac{dn}{d\tilde{w}}\bigg|_u + \tilde{w}\frac{dn}{dp}\bigg|_{\tilde{w}} \tag{9}$$

where we use the shorthand notation $\tilde{w}$ and p to denote the after-tax wage and nonlabor income, respectively. The term on the LHS is usually referred to as the "uncompensated" labor supply elasticity and the first term on the RHS as the compensated labor or "substitution" supply elasticity. The compensated elasticity is unambiguously positive (because the Hicksian demand for leisure unambigously slopes down) and the second term on the RHS is unambiguously negative (because leisure is a normal good).[10] In theory the uncompensated elasticity can be positive or negative; in the latter case the wealth effect dominates and labor supply is said to "bend backward."

To the extent that the "nonlabor income constant" perspective is relevant for empirical observations,[11] the uncompensated elasticity is more closely connected with the observed effects of taxes on behavior, while the compensated elasticity relates to the effect of taxes on efficiency. Here we sketch two approaches for obtaining estimates of the efficiency effects from observed behavior. The first approach uses the Slutsky equation, amended with a third term to capture "other determinants" of labor supply such as tastes or discrimination. In particular, consider a regression of log labor supply on log after-tax wage, nonlabor income, and other variables to capture the "other determinants." Assuming we appropriately measured the other determinants, or have satisfactorily dealt with their omission, the estimated coefficient on log after-tax wage is an estimate of the LHS of (9). The coefficient on nonlabor income, when multiplied by n, is an estimate of the second term on the RHS. We can then subtract the two to arrive at an estimate of the uncompensated elasticity.

A second approach is in the spirit of the Slutsky equation, but does not use it explicitly. Instead, the first step is to choose a parametric functional form for the utility or marginal rate of substitution function. For example, we might choose a CES utility function (and thereby a power MRS function), whose parameters are the elasticity of substitution in utility, and the relative weights on consumption and leisure. Second, we separately derive the behavioral and efficiency effects of taxes, as functions of the utility parameters. The third step is to "estimate" the utility parameters, by which I mean to pick those utility

---

[10] The various slopes in the Slutsky equation can be calculated in terms of the marginal rate of substitution function $M(c,l)$:

$$\frac{dn}{d\tilde{w}}\bigg|_s = \frac{1-nM_c}{\tilde{w}M_c - M_l}\,, \qquad \frac{dn}{d\tilde{w}}\bigg|_u = \frac{1}{\tilde{w}M_c - M_l} > 0\,, \qquad \frac{dn}{dp}\bigg|_{\tilde{w}} = \frac{-M_c}{\tilde{w}M_c - M_l} < 0$$

where M_c and M_l denote first partial derivatives. The final inequality follows from the assumption that leisure is a normal good.

[11] In our aggregate analysis, nonlabor income was not constant because market clearing and the government budget constraint dictate that higher tax rates affect transfer income and or the profit rates of firms. See also the caveat discussed above in connection with applications of the consumer theory perspective to the behavior of married women.

parameters that give the closest match for behavior in the model and behavior in the data.[12] The final step is to use the estimated utility parameter values, and the efficiency function from the second step, to calculate an exact value for the efficiency effect of taxes.

A detailed comparison of these two approaches is beyond the scope of this course, but I can offer two pieces of advice in reading the literature. First of all, the two approaches would be equivalent if in the second approach we choose exactly the right form for our utility function so that the Marshallian terms of the Slutsky equation were literally constants (rather than varying with consumption and leisure). In this sense, the two approaches differ only to the extent that the "right" utility function differs from this function. Second, the first approach mentioned above is older, and now sometimes considered old-fashioned. The second approach is often referenced in the modern literature as the "structural estimation" approach – a very confusing parlance because the distinction between the first and second approaches has absolutely nothing to do with the "structural" econometric models one might construct (using either approach) in order to appropriately account for the "other determinants" of labor supply.

IV. Advanced Topics

With a basic understanding of the economics of tax distortions, we have a good start on the study of a number of advanced topics, including:

- using the concept of "taxable income" to forecast Treasury revenues
- the importance of the behavior of rich people for tax collections
- the welfare economics of taxation
- applications of public finance to other fields
- the political economy of redistribution
- reasons for the growth of government

Advanced topics like these will capture our attention for the remainder of the course.

V. References

Barro, Robert J. and Chaipat Sahasakul. "Average Marginal Tax Rates from Social Security and the Individual Income Tax." *Journal of Business*. 59(4), October 1986: 555-66.

Gruber, Jonathan and David Wise, eds. *Social Security Programs and Retirement Around the World*. Chicago: University of Chicago Press for NBER, 1999.

Olson, Mancur. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press, 1965.

Trejo, Stephen J. "The Effects of Overtime Pay Regulation on Worker Compensation." *American Economic Review*. 81(4), September 1991: 719-40.

United States Congress, House Committee on Ways and Means. *Overview of Entitlement Programs, 1996 Green Book*. Washington, D.C.: U.S. Government Printing Office,

[12] In this third step we encounter essentially the same econometric issues as with the Slutsky equation approach – namely how to appropriately account for the "other determinants" of labor supply.

1996.

United States Social Security Administration. *Social Security Programs Throughout the World - 1995*. Washington, DC: Government Printing Office, 1995.

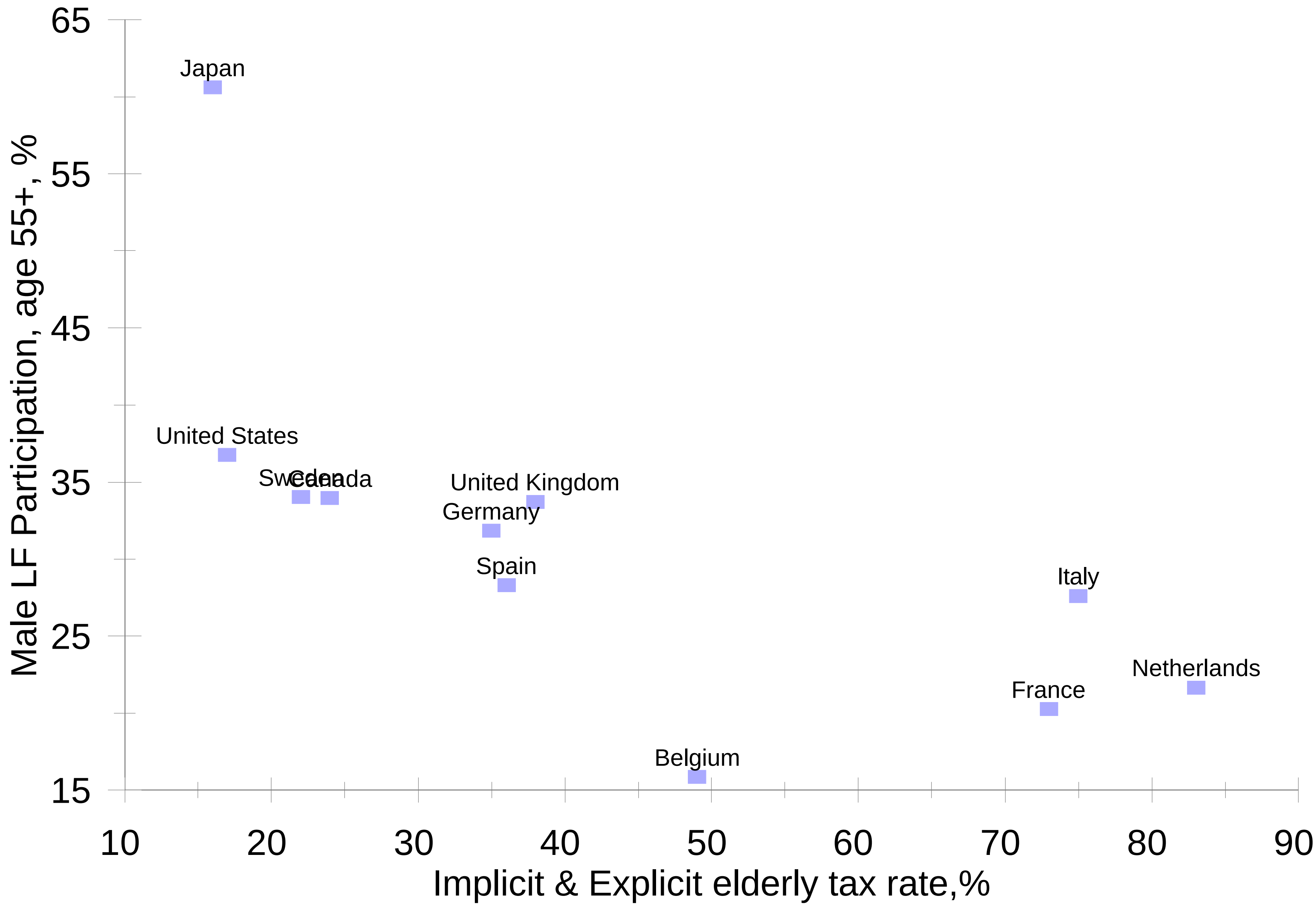

Chart 1. Elderly Work and Marginal Tax Rates, 11 countries