1 Excess burden of taxation

1. In a competitive economy without externalities (and with convex preferences and production technologies) we know from the 1. Welfare Theorem that there exists a decentralized equilibrium with prices that clears all markets and that is Pareto Efficient. The competitive price vector guarantees that all consumers have the same marginal rate of substitution between any pair of goods, which again is equal to the marginal rate of transformation between these goods on the production side. A deadweight loss arise if prices are distorted from this utopian benchmark. The deadweight loss measures the economic loss (in terms of lower consumer and producer surplus) caused by the price distortion. When it is taxes that distorts prices, that is, when it is a tax that drives a wedge between consumer and supplier prices, we talk about the excess burden of taxation (or the deadweight loss associated with taxation). The excess burden of a tax (or of a tax system) is the economic loss tax payers experience, over and above the tax revenue that is collected by the government. If consumers experience a loss, measured in NOK, of magnitude I when a tax is introduced and the revenue collected is R, then the deadweight loss is $I - R$. There is usually an excess burden associated with taxation since taxes do distort prices in such a way that valuable transactions do not happen. A tax on labour income will distort the price of leisure and induce a dead weight loss since the marginal rate of substitution between consumption and leisure is different from the marginal productivity of labour. In some cases – mostly in theory – a tax is levied on a tax base that is exogenously given. In this case there is no excess burden of taxation. In other cases, and this is a more relevant situation, taxes can correct prices that are set to low in the market because they do not include the costs of negative externalities.

2. One definition of the deadweight loss of a tax is the amount the government would have to give the consumer to compensate for the taxes, net of the revenue the it collects. Let prices before taxes be fixed and equal to $\mathbf{q_0}$, after price tax is given by $\mathbf{q_1} = \mathbf{q_0} + \mathbf{t}$. Let $V(\mathbf{q}, y)$ be the representative households indirect utility function. One measure of the excess burden

$$EB^{cv} = E(\mathbf{q_1}, V(\mathbf{q_0}, y) - y - \mathbf{t}'\mathbf{h}(\mathbf{q_1}, V(\mathbf{q_0}, y))$$

   This is the compensating variation measure of excess burden ($h()$ is the compensated demand function) . If we instead use the after tax utility level as a reference we obtain the equivalent variation measure

$$EB^{cv} = y - E(\mathbf{q_0}, V(\mathbf{q_1}, y) - \mathbf{t}'\mathbf{x}(\mathbf{q_1}, y)$$

3

(a) There is no EB associated with this tax. The EB is the familiar Harberger triangle and this vanishes as the tax that is introduced goes towards zero. The point is that for a very small tax the value lost for the consumer is exactly equal to the revenue collected by the government. If the tax on this good was increased further there would be an EB and this loss will increase more than proportionally with the tax increment.

(b) In this case the tax on i has an impact on the EB by changing the tax revenue for the government since the price change on good i will have impact on the demand of other goods.

2 Public Goods and Taxation

The utility function is fully separable in consumption of the different goods and it is linear in the consumption of good x and concave in the consumption of y and the public good g. This means that x absorbs all the income effects. Note also that the log specification implies that the gross expenditures on good y is independent of the price on good y and is equal to unity. The production technology is linear, one has to give up one unit of a private good in order to produce one unit of the public good.

1. With lump sum taxes the problem is to maximize (since all individuals are equal) $n{*}u(x, y, g)$ subject to the constraint that $nw = nx+ny+g$. We obtain the necessary contition that $g^* = \alpha n$. That is, the government should provide the public good until $n \cdot MRS_{gx} = MRT_{gx} \Rightarrow n \cdot \left(\frac{\partial U}{\partial g}/\frac{\partial U}{\partial x}\right) = 1$: the Samuelson-rule.

2. If the public good must be financed by a per unit tax on good y the budget constraint for a household is $x+(1+t)y = w$. Given the utility function maximal utility implies $y = \frac{1}{1+t}$ and $x = w-1$. The Government maximizes $n\left[(w-1) - ln(1+t) + \alpha ln(g)\right]$ subject to the constraint that $g = n \cdot t \cdot \frac{1}{1+t}$. Solving this problem we obtain $t = \alpha$ and hence that $g^{**} = \frac{\alpha n}{1+\alpha} < g^*$. Which is pretty obvious: when the government has to finance the public good with a distortionary tax (and there is no other distortion around) this implies less public good and the government should stop before the Samuelson-Rule.

3. Taxation can, in some cases, imply that the government should produce public goods beyond the level where $\sum MRS_{gx} = MRT_{gx}$. We have discussed two possibilities (with an emphasis on the first)

   (a) If the government cannot redistribute as much money as it wants to the needy because the self selection constraint for the not-so-needy binds it can be optimal to over-provide public good to relax this constraint. This is the case if the not-so-needy has a lower MRS_{gx} than the needy if they choose to mimic the needy.

(b) In a Ramsey set-up it can be shown that if the public good is complementary to the non-taxed good (leisure) it can be optimal to overprovide the good.

Guide to answering Problem 4, Econ 4620, Spring 2015

The Elasticity of Taxable Income (ETI) and the discrete choice labor supply model refer to two different methods to obtain information about individual tax responsiveness. Present the two methods and discuss their advantages.

Answers to Problem 4 are found "Validation of the Discrete Choice Labor Supply Model by Methods of the New Tax Responsiveness Literature" (Thoresen/Vattø).

Regarding the ETI approach, the student can give a less formal description, which should include some of the following: [1]

In this perspective, the so-called elasticity of taxable income approach (the ETI approach), or interchangeably labelled the new tax responsiveness literature (the NTR approach), represents a promising alternative for use in external validations, as it denotes a well-established procedure to rinse out the effects of taxes. Studies of the large and growing ETI/NTR literature exploit that tax reforms generate net-of-tax rate changes along the income scale, often resulting in substantial tax changes for some tax-payers, whereas others are more or less unaffected. Taxable income is used as the main measure of outcome in this literature, as it in principle captures all the public policy relevant behavioral responses of a reform (hours worked, effort, tax avoidance and evasion, change of job, etc).

The student should also mention that income panel data are often used in the estimation. A good score should explain more precisely, less formally or more formally, how tax effects are identified, along the following lines (page 9, 10):

Panel data covering a period of net-of-tax rate variation across individuals and across time (often covering a tax reform) have been the main data source for the identification of responses in the empirical framework of the NTR approach. Taxable income for individual i at time t, q_{it}*, is explained by a time-specific constant,* κ_t*, the net-of-tax rate,* $\log(1-\tau_{it})$*, unobserved heterogeneity* μ_i *and the remaining iid error term,* ξ_{it}*,*

$$\text{(2.2)} \qquad \log q_{it} = \kappa_t + \lambda \log(1-\tau_{it}) + \mu_i + \xi_{it} .$$

The basic framework for identification in the NTR literature consists of various estimations of a first-differenced version of (2.2), using panel data for two periods,

$$\text{(2.3)} \qquad \Delta \log q_i = \kappa + \lambda \Delta \log(1-\tau_i) + \Delta \xi_i .$$

[1] Note that the ETI (elasticity of taxable income) and the new tax responsiveness literature (NTR) refer to the same concept.

The coefficient of interest, λ, measures the elasticity of income with respect to changes in the net-of-tax rate defined as $\frac{1-\tau}{q}\frac{\partial q}{\partial(1-\tau)}$. *The reliability of results depends on carefully framed empirical designs for the identification of the key parameter, including controls for individual characteristics that might affect income growth. One obvious methodological identification challenge (w.r.t. λ) has been the endogeneity of the tax rate, which has led to the estimation of (2.3) using IV techniques. For instance, Feldstein (1995) employs the difference-in-differences estimator, and let the change in the net-of-tax rates and the allocation into groups (groups more or less treated by the US tax reform of 1986) be determined by pre-reform income levels. Many post-Feldstein studies employ a closely related exclusion restriction, namely the change in net-of-tax rates based on a fixed first period income as instrument in an IV regression; see Auten and Carroll (1999) and Gruber and Saez (2002). Thus, the NTR literature is related to methods commonly used in the "experimentalist" or "program evaluation" literature. However, the conventional identification technique of the NTR literature implies that one is far from an ideal randomized trial situation.*

An application of the discrete choice labor supply model to derive information about tax responsiveness should emphasize that now one estimates a model based on utility maximizing, and the model is in turn used to simulate the effects of tax changes (page 1)

Based on cross-sectional observations of households' and individuals' consumption and connections to the labor market (typically working hours), labor supply models can be estimated and then used in the policy-making process for simulations of short term labor market effects of prospective changes in the tax system.

A good is obtained if some of the following details are provided: further description of the utility function, about the estimation procedure (conditional logit ala McFadden), further specification of the utility (as Box-Cox), as on page 6:

With the discrete choice approach, it is easy to deal with nonlinear and nonconvex economic budget constraints, and to apply rather general functional forms of the utility representation.

With particular distributional assumptions about the stochastic disturbances in the utility function one can derive tractable expressions for the distribution of hours of work, such as the multinomial logit model or the nested multinomial logit model. The maximization problem for a person in a single-individual household can be seen as choosing between bundles of consumption (C) and leisure (L), subject to a budget constraint, $C = f(hw, I)$, *where h is hours of work, w is the wage rate, I is non-labor income, C is (real) disposable income and f(·) is the function that transforms gross income into after-tax household income.*

The utility function of the household is assumed to be additively separable, $U(C,h) = \nu(C,h) + \varepsilon(C,h)$, *where* $\nu(\cdot)$ *is a positive deterministic function and* ε *the random unobserved components for individual i and choice j. We assume that the random components are i.i.d. extreme value distributed with c.d.f.* $\exp(\exp(-x))$ *for positive x, which implies independence of* *irrelevant alternatives (IIA). The strict IIA assumption can be weakened, however, by allowing for random effects in utility parameters or in relation to the wage rate.*[2]

Let $v(\cdot)$ be the representative utility of jobs with hours of work h, a given individual specific wage rate w, and non-labor income I. By applying standard results in discrete choice theory (McFadden, 1984), it follows that the probability that the agent will choose working hours h can be expressed as

$$(2.1) \qquad P(h)=\frac{\exp v(f(hw,I),h)}{\exp v(f(hw,I),0)+\sum_{h\in D}\exp v(f(hw,I),h)}.$$

We see different specifications of the deterministic part in the literature.[3] *Here, we use a flexible Box-Cox functional form specification,* $v(C,h)=\alpha_0\frac{(C-C_0)^{\alpha_1}-1}{\alpha_1}+(\beta_0+\gamma Z)\frac{(\bar{h}-h)^{\beta_1}-1}{\beta_1}$*, where C measures the household-adjusted consumption level, constructed by dividing the couple or individual's disposable income by* $\sqrt{N}$*, where N is the number of individuals in the household (including children under 18).*

A discussion of the advantages of the two methods should include that the ETI approach is simpler and (perhaps) relies on less assumptions in deriving results. But that some sort of structural modeling (as the discrete choice model is needed in order to simulate results of prospective policies. In addition, in the field of structural modeling, the discrete choice variant is popular because it is simpler than the alternatives. Further, the criticism of the "structural approach", as being too stylized and may suffer from misspecification, can be seen be used as arguments for using the ETI framework. Discussion of data sources and data availability should be credited.

---

[2] We replace the wage rate by a wage equation that includes a stochastic error term, and thus a mixed multinomial logit model follows, see McFadden and Train (2000) and Haan (2006).

[3] Quadratic or translog functional forms for the systematic part of the utility function have also been used in several applications, see, e.g., van Soest (1995). One advantage of the Box-Cox functional form is that it is globally monotone in consumption and leisure; see Dagsvik et al. (2014) for a discussion of this issue. In practice, the choice of functional form seems to have little impact on results.