Mixed Taxation

Laurence Jacquet

THEMA and OFS

September 2016

Outline

Content

Taxes on goods and linear labor income:
Ramsey problem, Corlett-Hague Theorem (ReStud 1953) and generalization by Harberger (1964)

Content

Taxes on goods and nonlinear labor income:

- Atkinson-Stiglitz Theorem
- Atkinson-Stiglitz revisited by Konishi (JPubE 1995) Laroque (EL 2005) and Kaplow (JPubE 2006)
- Extensions of Atkinson-Stiglitz theorem: heterogeneous needs, heterogeneous tastes or preferences, emission of an externality by one or more goods, imperfect substituability of different skills of labor, wage uncertainty, dynamic optimal taxation

- On this topic, I strongly advise students to read **Chapter 3 in** the excellent book of **Boadway (2012), pp.47-84**, before reading the literature
- All references can be found at the end of this handout.

The Ramsey rule gives the optimal (linear) tax rates on commodities when the **labor income tax** is constrained to be **linear**.

Optimal Commodity Taxation: Ramsey Problem

Representative individual benchmark case

- Utility: $u(x_1, ..., x_n, h - \ell) = u(x_1, ..., x_n, x_0)$ where h: total time the individual has available
- Production has CRS of the simplest variety: each good is produced from labor alone.

Production of a unit of good i requires a_i units of labor so that the production price can only be $p_i = a_i w$ where w is the wage per unit of labor in equilibrium.
Normalize $w = 1$; choose the units of goods so that each a_i equals one, so that all production prices satisfy $p_i = 1$ $(i = 1, \cdots, n)$.

- **Linear** taxes on goods and **linear** tax on wage

- Consumer prices are tax inclusive: $q_i = 1 + t_i$, $i = 1, ..., n$.
- Price of leisure or labor is $p_0 = w = 1$, after-tax wage: $(1 - \widetilde{t})$ and we will see that we can assume $\widetilde{t} = 0$; $(t_0 = 0)$.
- The budget constraint of consumer (who only owns her own labor force) then is

$$\sum_{i=1}^{n} q_i x_i = (1 - \widetilde{t})\ell \tag{1}$$

In this setting with no nonlabor income, and no bequests, **the tax on wage is equivalent to a uniform tax on goods ⇒ tax on wage can be neglected.** Indeed define

$$t_i' = \frac{\widetilde{t} + t_i}{1 - \widetilde{t}}.$$

Since $1 + t_i' = \frac{1+t_i}{1-\widetilde{t}} = \frac{q_i}{1-\widetilde{t}}$, we can rewrite the individual budget constraint, defining $q_i' = 1 + t_i'$, as

$$\sum_{i=1}^{n} q_i' x_i = \ell. \tag{2}$$

The tax system $\left((t_i), \widetilde{t} \right)$ is then **equivalent** to the tax system $\left((t_i'), 0 \right)$ which does not tax wages. Replacing $\left((t_i), \widetilde{t} \right)$ with $\left((t_i'), 0 \right)$ leaves consumer choices unchanged.

Moreover, the gvt collects from one consumer with $\left(\left(t_i\right), \widetilde{t}\right)$:

$$\sum_{i=1}^{n} q_i x_i + \widetilde{t}\ell,$$

using the individual budget constraint $\ell = \sum_{i=1}^{n} q_i' x_i$, this tax revenue can be rewritten as

$$\sum_{i=1}^{n} \left(t_i + \widetilde{t}\left(1 + t_i'\right)\right) x_i = \sum_{i=1}^{n} q_i' x_i$$

which is exactly what the gvt collects from this consumer with the tax system $\left(\left(t_i'\right), 0\right)$. $\Rightarrow$ **A (linear) tax on wages is absolutely equivalent to a uniform tax on goods**.

- $\Rightarrow$ Tax on labor is then suppressed because it is equivalent to a uniform tax on goods x_i and is therefore redundant. From now on: $\widetilde{t} = 0$.
- $\Rightarrow$ Only n of the $(n+1)$ tax rates are determined at the optimum, whatever that is.

$\Rightarrow$

Representative individual problem:

$$\underset{x_i,\ell}{Max}\ u(x_1, ..., x_n, h-\ell)$$

$$\text{s.to: } \sum_{i=1}^{n} q_i x_i - \ell = 0 \text{ (mult. } \alpha)$$

$\Rightarrow \mathbf{x}(\mathbf{q})$, $\ell(\mathbf{q})$ and $v(\mathbf{q}) \equiv u\left((\mathbf{q}),\ell(\mathbf{q})\right)$ is the indirect utility

Government problem

$$\underset{x_i,\ell}{Max}\; v(\mathbf{q}) \text{ s.to: } \sum_{i=1}^{n} t_i x_i(\mathbf{q}) = R \text{ (mult. } \lambda)$$

$$\mathcal{L}(\mathbf{t},\lambda) \equiv v(q) + \lambda \left[\sum_{i=1}^{n} t_i x_i\left(\mathbf{q}\right) - R \right]$$

First-Order Conditions:

FOC (t_k):

$$\underbrace{\frac{\partial v}{\partial q_k}}_{\text{impact on utility}} \underbrace{+\lambda \left[x_k + \sum_{i=1}^{n} t_i \frac{\partial x_i}{\partial q_k} \right]}_{\text{impact on tax revenue}} = 0 \quad k = 1, ..., n$$

FOC (λ):

$$\sum_{i=1}^{n} t_i x_i\left(\mathbf{q}\right) - R = 0 \quad k = 1, 2$$

The solution to these conditions yield $t_k\left(R\right)$ and $\lambda\left(R\right)$.

The Ramsey rule

Envelope theorem: $\frac{\partial v}{\partial q_k} = -\alpha x_k(\mathbf{q}, w)$ $(k = 1, ..., n)$ i.e. Roy's identity (The derivative of the indirect utility function w.r.to the price of good k is equal to the marginal utility of income times the quantity of good k).
Hence FOC (t_k) becomes:

$$\underbrace{-\alpha x_k}_{\partial v/\partial q_k} + \lambda \left[x_k + \sum_{i=1}^{n} t_i \frac{\partial x_i}{\partial q_k} \right] = 0 \quad k = 1, ..., n$$

$$\Leftrightarrow \sum_{i=1}^{n} t_i \frac{\partial x_i}{\partial q_k} = -\frac{\lambda - \alpha}{\lambda} x_k \quad k = 1, ..., n$$

The Ramsey rule

Remember the Slutsky equation:

$$\frac{\partial x_i}{\partial q_k} = \underbrace{S_{ik}}_{\substack{\text{substitution}\\ \text{effect}}} - x_k \frac{\partial x_i}{\partial I}$$

with

$$I : \text{a lump-sum income and}$$

$$S_{ik} = \left.\frac{\partial x_i}{\partial q_k}\right|_u$$

Rewrite the FOC (t_k) using the Slutsky equation:

$$\frac{\sum_i t_i S_{ik}}{x_k} = -\frac{\lambda - \alpha}{\lambda} + \sum_i t_i \frac{\partial x_i}{\partial I} \quad k = 1, \ldots, n$$

The Ramsey rule

$$\frac{\sum_i t_i S_{ik}}{x_k} = -\frac{\lambda - \alpha}{\lambda} + \sum_i t_i \frac{\partial x_i}{\partial I} \quad k = 1, ..., n$$

Note the RHS is the same for all equations, i.e. for all goods. We can then rewrite the **Ramsey** proportionate reduction **rule** as:

$$\sum_i t_i S_{ki} = -\theta x_k \qquad k = 1, ..., n \qquad \text{(Ramsey R.)}$$

with

$$\theta = 1 \quad - \underbrace{\left[\frac{\alpha}{\lambda} + \sum_i t_i \frac{\partial x_i}{\partial I}\right]}_{\text{net social mg utility of income}}$$

where we have used the symmetry property of Slutsky substitution matrix: $S_{ik} = S_{ki}$.

Remark: **The value of θ is independent of the particular good chosen.**

The Ramsey rule

sign of θ?

Multiply (Ramsey R.) by t_k and sum on k:

$$\underbrace{\sum_k \sum_i t_i S_{ik} t_k}_{\substack{<0 \text{ because the (Slutsky)} \\ \text{matrix of } S_{ik} \text{ is negative} \\ \text{semi-definite}}} = -\theta \sum_k t_k x_k = -\theta R$$

$\Rightarrow \theta R > 0$ (θ and R have same sign)

Interpretation: One should design the tax system so as to induce the same total proportional distortion on all the goods

Justification of this interpretation:
Let $dt_i = t_i d\beta$ be the small **proportional** rise **of all taxes** $(i = 1, ..., n)$, from the Ramsey equation:

$$\underbrace{\sum_i S_{ki} t_i d\beta}_{dx_k} = -\theta x_k d\beta \qquad k = 1, ..., n$$

$\sum_i S_{ik} t_i$ is an approximation to the total change in **compensated** demand for good k due to the taxation change
Therefore:

$$\frac{\left.\frac{dx_k}{d\beta}\right|_u}{x_k} = -\theta \qquad k = 1, ..., n$$

Interpretation: One should design the tax system so as to induce the same total proportional distortion on all the goods

$$\frac{\left.\frac{dx_k}{d\beta}\right|_u}{x_k} = \underbrace{-\theta}_{\text{constant}} \qquad k = 1, \ldots, n$$

Optimal commodity taxes require that a small **proportional** rise of all taxes reduces the compensated demand for all goods **in the same proportion**.
($\simeq$ interpretation with inverse elasticity rule!)

Other interpretation of Ramsey rule: tax system optimal when equal discouragement index

$$\frac{\sum_i t_i S_{ik}}{x_k} = -\theta \qquad k = 1, ..., n \qquad \text{(Ramsey R.)}$$

Ramsey rule says that the tax system is optimal **when the index of discouragement is equal for all goods** (Mirrlees JPubE 1976)**:**

$$d_k = \frac{\sum_i t_i S_{ik}}{x_k} \qquad k = 1, ..., n$$

Index of discouragement of good $k \stackrel{\text{def}}{\equiv}$ proportional reduction in demand of good k

Ramsey rule:

$$d_k = -\theta \qquad k = 1, ..., n$$

Further insight into the Ramsey Rule by considering the case of two goods and leisure

- Corlett and Hague Theorem (ReStud, 1953): For a representative consumer choosing three commodities (2 goods and leisure), revenue-neutral deviations from a uniform tax on the two goods would improve the welfare of the individual if a higher tax rate were imposed on the good more complementary with leisure.
- Harberger (1964) generalizes this three-commodity case to the choice of an optimal commodity tax system; it is optimal that goods that have a higher complementarity with leisure have a higher ad valorem tax rate (t_k / q_k).

Example:
What about urban transportation?
More complementary with labor than with leisure $\Rightarrow$ should be taxed less heavily

The following slides give a proof. Rewriting Ramsey rule in the three-commodity case gives:

$$S_{11}t_1 + S_{12}t_2 = -\theta x_1$$

$$S_{21}t_1 + S_{22}t_2 = -\theta x_2$$

Cramer's rule: with $S = S_{11}S_{22} - S_{12}S_{21}$

$$t_1 = [S_{12}x_2 - S_{22}x_1]\theta S^{-1}$$

$$t_2 = [S_{21}x_1 - S_{11}x_2]\theta S^{-1}$$

and so

$$\frac{t_1/q_1}{t_2/q_2} = \frac{q_2 S_{21}/x_1 - q_2 S_{22}/x_2}{q_1 S_{21}/x_2 - q_1 S_{11}/x_1}$$
$$= \frac{\varepsilon_{12}^c - \varepsilon_{22}^c}{\varepsilon_{21}^c - \varepsilon_{11}^c} = \frac{\varepsilon_{22}^c - \varepsilon_{12}^c}{\varepsilon_{11}^c - \varepsilon_{21}^c} \qquad \text{(A)}$$

with $\varepsilon_{ij}^C = \frac{S_{ij} q_j}{x_i} = \frac{q_j}{x_i} \frac{\partial x_i}{\partial q_j}\Big|_u$: compensated elasticity of demand for a good i with respect to price j.

Using **homogeneity of compensated demands**, $\sum_{i=0}^{n} \varepsilon_{ji}^{c} = 0$ and defining the ad valorem tax rate with $\tau_i = t_i / q_i$, we obtain:

$$\frac{\tau_1}{\tau_2} = \frac{\varepsilon_{22}^{c} + \varepsilon_{11}^{c} + \varepsilon_{10}^{c}}{\varepsilon_{11}^{c} + \varepsilon_{22}^{c} + \varepsilon_{20}^{c}}$$

Since $\varepsilon_{11}^{c} + \varepsilon_{22}^{c} < 0$, $\tau_1 > \tau_2$ iff $\varepsilon_{10}^{c} < \varepsilon_{20}^{c}$ i.e., if x_1 is relatively more complementary with leisure than x_2

The more complementary is a good with leisure the higher will be its ad valorem tax rate (Harberger 1964).
Intuiton: Since leisure is untaxable, imposing a higher tax on the good that is more complementary with leisure is an indirect way of taxing it.

Remark: “a good more complementary with leisure" means:
For a given level of utility, a rise in the wage rate (which increases labor supply) reduces the consumption of this good.

Implication of these results: If $\varepsilon_{10} = \varepsilon_{20}$, so goods x_1 and x_2 are equally complementary with leisure, **welfare cannot be improved by deviating from uniformity of commodity tax rates**. This result readily generalizes to any number of goods and leisure:

$$\tau_i = \tau \text{ if } \varepsilon_{k0} = \varepsilon_{j0} \quad \forall j, k = 1, ..., n.$$

This result on **uniform commodity taxation = Corlett-Hague theorem** (Corlett and Hague Restud, 1953).

- A utility function satisfying $\varepsilon_{k0} = \varepsilon_{j0} \quad \forall j, k = 1, ..., n$ is the weakly separable one of the form $u\left(f\left(x_1, ..., x_n\right), x_0\right)$ with the **sub-utility** function f being **homothetic** (Sandmo JPubE, 1976).
- Intuitively, preferences for goods are independent of leisure. An increase on tax w will therefore decrease labor income available for spending on goods, and since **preferences are homothetic**, the demand for all goods will decrease proportionately. A proportional tax on all goods will be equivalent to a tax on leisure and will cause the demand to fall proportionately thereby not imposing any distortion in demand patterns.

Outline

Mixed taxation: Atkison-Stiglitz

Until now: linear labor income tax
In practice, the labor income tax is nonlinear

Atkison-Stiglitz theorem: When **labor income** is (optimally) **taxed nonlinearly**, if utility is weakly separable in leisure and goods so that the utility takes the form $u\left(f\left(x_1, ..., x_n\right), x_0\right)$, optimal commodity tax should be uniform **regardless of the properties of the sub-utility function in goods** (Atkinson and Stiglitz JPubE, 1976).

Note:

- Prefences satisfy taste homogeneity
- This result is not whether to tax labor income or consumption. It is possible to reduce the average tax rate on labor income and increase that on consumption without changing the allocation of resources.The question is which commodities, if any, should be taxed more than others.

Atkison-Stiglitz proof in the context of Mirrlees (ReStud 1971)

Atkison and Stiglitz (1980), Christiansen (JPubE 1984), Boadway (2012, pp.60-61)

Commodities:
n commodities indexed by i.
Each commodity i faces:

- Goods x_i , $i = 1, ..., n$
- Commodity taxes t_i; normalize $t_1 = 0$
- Production prices are 1 (see before)
- A consumption price $q_i = 1 + t_i$ where t_i is the tax rate
- Remark: nonlinear commodity taxes are not excluded in Atkinson, Stiglitz (JPubE 1976), i.e. t_i can be a function of the level of consumption of good x_i

Consumers:

There is a continuum of consumers. Each Consumer has:

- Productivity w (which is private information), $w \in [\underline{w}, \overline{w}]$ where $\underline{w} \geq 0$, $\overline{w} \leq \infty$.
- Labor supply ℓ (which is also private information)
- Consumption x_i of commodity i.

Type-w individual problem:

$$\underset{x_i, \ell}{Max}\ u\left(x_1, ..., x_n, \ell\right)$$

$$\text{s.to: } \sum_{i=1}^{n} \left(1 + t_i\right) x_i = w\ell - T\left(w\ell\right) \qquad \text{(mult. } \alpha\text{)}$$

where $\ell\left(w\right) = y\left(w\right)/w$ is labor supply and $T(w\ell)$ is the non-linear labor income tax schedule.

Individual First-Order Conditions:

$$\begin{aligned}\frac{u_{x_k}}{u_{x_j}} &= \frac{1+t_k}{1+t_j} = \frac{q_k}{q_j},\\ \frac{wu_{x_k}}{u_\ell} &= -\frac{1+t_k}{1-T'} = -\frac{q_k}{1-T'}\end{aligned}$$

Government problem (Mirrlees)

$Max \int W(u) f(w)\, dw$

s. to

$$\int \left(w\ell - \sum x_i - R \right) f(w) dw \qquad \text{(mult. } \lambda\text{)}$$

where R is revenue requirements and s. to incentive compatibility:

$$\dot{u} = -\ell u_\ell / w \qquad \text{(mult. } \zeta\text{)}$$

which is a sufficient condition for the incentive constraints $\forall (w, w') : u(\mathbf{x}(w), y(w)/w) \geq u(\mathbf{x}(w'), y(w')/w)$.

Control variables: $x_2, ..., x_n, \ell$ with x_1 determined by inverting the utility function, $x_1 = (x_2, ..., x_n, \ell, u)$.
State variable u (all variables vary continuously with w)

Remark quantities of taxed goods are control variables so that Gvt can employ nonlinear commodity taxes (see later)

Hamiltonian:

$$H = \left(W\left(u\right) + \lambda\left(w\ell - \sum x_i - R\right)\right) f\left(w\right) dw - \zeta l u_\ell / w$$

Interested in indirect taxes: First-Order Condition on x_k for a type-w individual:

$$-\lambda\left(1 + \left.\frac{\partial x_1}{\partial x_k}\right|_u\right) f\left(w\right) - \zeta\frac{\ell}{w}\left(u_{\ell k} + u_{\ell x_1}\left.\frac{\partial x_1}{\partial x_k}\right|_u\right) = 0$$

Using $\left.\frac{\partial x_1}{\partial x_k}\right|_u = \frac{-u_{x_k}}{u_\ell} = -\left(1 + t_k\right)$, this can be written

$$t_k = \frac{\zeta_\ell u_k}{\lambda wf\left(w\right)}\left(\frac{u_{\ell x_k}}{u_{x_k}} - \frac{u_{\ell x_1}}{u_1}\right), \text{ or } \frac{t_k}{q_k} = \frac{\zeta \ell \alpha}{\lambda wf}\left(\frac{d \log\left(u_{x_k} / u_{x_1}\right)}{d\ell}\right),$$

where α is the consumer's marginal utility of income. $\Longrightarrow$

- If utility takes the **separable utility form** $u\left(f\left(\mathbf{x}\right), \ell\right)$, $\frac{t_k}{q_k} = 0$ for all $k = 1, ..., n$, i.e. uniform commodity taxation, which is the **Atkinson-Stiglitz (A-S) theorem**.
- $\frac{t_k}{q_k} > 0$ iff x_k **is more complementary with leisure** than x_1 in the sense that **its relative valuation increases with labor** supplied. This bears an **uncanny resemblance to** the **Corlett-Hague** theorem in the three-commodity case.

Remark: “a good more complementary with leisure" here means:
For a given level of income, a rise in labor supply reduces the consumption of this good.

Recall in Corlett-Hague and Harberger, it meant: For a given level of utility, a rise in the wage rate (which increases labor supply) reduces the consumption of this good.

Remarks:

- The expression above is derived for a type-w individual, optimal t_k differs for each type **if weak separability does not apply**.
- If **no** weak-separability, the optimal tax structure can then be implemented if **nonlinear commodity taxes** are deployed (Atkinson and Stiglitz JPubE, 1976).
- If one restricts to the **same linear commodity tax rates for all individuals**: **higher taxes on goods that are more complementary with leisure**.

Atkison-Stiglitz proof in the two types case

Edwards, Keen and Tuomala (FinanzArchiv, 1984), Nava, Schroyen and Marchand (JPubE, 1996),
Boadway (2012, pp.61-77)

Assumptions

- 2 types of productivity w^j $(j = 1, 2)$, with $w^2 > w^1$.
- n^j = number of individuals of type j.
- Goods x_1 and x_2, and labour ℓ.
- Utility function: $u(x_1, x_2, \ell)$.

- Gross labor income of type-j : $y^j = w^j \ell^j$.
- The Gvt can impose a nonlinear income tax on y as well as direct commodity taxes on goods.
- Normalize tax on x_1 to zero so that the only indirect tax will be t_2 on x_2.
- Assume producer prices are unity so that $q_1 = 1$ and $q_2 = 1 + t_2$

Individual Decisions

Disaggregate individual decision-making into **two stages**, given the tax system in place:

1. Choice of j to determine y^j and disposable income $c^j = y^j - T(y^j)$.
2. Allocation of c^j between x_1^j and x_2^j.

Stage 2 (where disposable and gross incomes are given)**:** Choice of x_1 , x_2.
Utility $u^j\left(x_1, x_2, y^j / w^j\right)$
Type-j 's problem, given $c^j = x_1^j + qx_2^j$ from Stage 1:

$$\underset{x_2^j}{Max}\, u^j\left(c^j - qx_2^j, x_2^j, y^j / w^j\right), \;\; j = 1, 2$$

$\Rightarrow x_2^j\left(q_2, c^j, y^j\right)$ which is decreasing in q_2. Let the value function for this stage be $v^j\left(q_2, c^j, y^j\right)$, with the associated Envelope theorem properties: $v_q^j = -x_2^j u_x^j$, $v_c^j = u_x^j$, $v_y^j = u_y^j / w^j$.

Mimicker's problem

A type-2 person might have been tempted to mimic a type-1 in stage 1. For a type-2 mimicking a type-1 (use a hat for mimicking), problem is:

$$\underset{\widehat{x}_2^j}{Max} \widehat{v}^2 \left(c^1 - q\widehat{x}_2^j, \widehat{x}_2^j, y^1 / w^1 \right), \quad j = 1, 2$$

$\Rightarrow$ we find $\widehat{x}_2^j \left(q_2, c^1, y^1 \right)$ and $\widehat{v}^2 \left(q_2, c^1, y^1 \right)$.the value function

Note: Type 1 and mimicker have same c^1, y^1, so $\ell^1 > \widehat{\ell}^2$ since type-2's have a higher wage rate.

If $u \left(x_1, x_2, \ell \right) = u \left(f \left(x_1, x_2 \right), \ell \right)$, then $x_2^1 = \widehat{x}_2^2$.

If instead x_2 is **more complementary with leisure** than is x_1, then $x_2^1 < \widehat{x}_2^2$.

Stage 1: Choice of c^j, y^j, Government Policy

- In stage 1, individuals choose labor supply, and therefore income. Following Stiglitz (1982), this choice can be treated as the outcome of a direct mechanism whereby the Gvt chooses $\left(c^j, y^j\right)$ for each of the types, such that these choices maximize social welfare while being incentive compatible and satisfying a revenue constraint.
- The Gvt chooses also the indirect tax rate t_2 (recall that $t_1 = 0$ by normalization)
- Disaggregate government policy into two stages:

1. Choice of nonlinear tax, given t_2
2. Choice of t_2

- Utilitarian Gvt

$$\underset{c^j, y^j}{Max} \;\; nv^1\left(q_2, c^1, y^1\right) + n^2 v^2\left(q_2, c^2, y^2\right)$$

$$\text{s.to: } n^1\left(y^1 - c^1 + t_2 x_2^1\left(q_2, c_1, y_1\right)\right)$$
$$+ n^2\left(y^2 - c^2 + t_2 x_2^2\left(q_2, c_2, y_2\right)\right) = R \text{ (mult. } \lambda\text{)}$$

and s.to (binding) incentive constraint:

$$v^2\left(q_2, c^2, y^2\right) = \widehat{v}^2\left(q_2, c^1, y^1\right) . \text{(mult. } \gamma\text{)}$$

Assume IC applies to type-2's.

Stage 1: Choice of c^j, y^j, Government Policy
Government Lagrangian:

$$L = \underset{c^j, y^j}{Max} \;\; n^1 v^1 \left(q_2, c^1, y^1\right) + n^2 v^2 \left(q_2, c^2, y^2\right)$$
$$+\gamma \left[v^2 \left(q_2, c^2, y^2\right) - \widehat{v}^2 \left(q_2, c^1, y^1\right)\right]$$
$$+\lambda \left[n^1 \left(y^1 - c^1 + t_2 x_2^1 \left(q_2, c_1, y_1\right)\right) + n^2 \left(y^2 - c^2 + t_2 x_2^2 \left(q_2, c_2, y_2\right)\right)\right] =$$

The FOC on t_2 reduces after some manipulations (by Envelope theorem and using FOC's w.r.to c^j and y^j) to:

$$\gamma\widehat{v}^2\left(\widehat{x}_2^2 - x_2^1\right) + \lambda t_2\left(n^1\frac{\partial\widetilde{x}_2^1}{\partial q_2} + n^2\frac{\partial\widetilde{x}_2^2}{\partial q_2}\right) = 0$$

where $\widetilde{x}_2^j$ is compensated demand for x_2 by person j.

Therefore, the optimal tax rate t_2^* can be written

$$t_2^* = -\frac{\gamma \widehat{v}_x^2 \left(\widehat{x}_2^2 - x_2^1 \right)}{\lambda \left(n^1 \partial \widetilde{x}_2^1 / \partial q_2 + n^2 \partial \widetilde{x}_2^1 / \partial q_2 \right)} \gtrless 0 \text{ as } \widehat{x}_2^2 \gtrless x_2^1$$

Denominator: Efficiency effect showing that t_2^* will be higher the smaller are the compensated demand responses to t_2 changes.

$$t_2^* = -\frac{\gamma \widehat{v}_x^2 \left(\widehat{x}_2^2 - x_2^1\right)}{\lambda \left(n^1 \partial \widetilde{x}_2^1 / \partial q_2 + n^2 \partial \widetilde{x}_2^1 / \partial q_2\right)} \gtrless 0 \text{ as } \widehat{x}_2^2 \gtrless x_2^1$$

Numerator:

- It involves the ICC. It will be zero if the utility function is $u\left(f\left(x_1, x_2\right), \ell\right)$ so mimicker demands same amount of x_2 as does the type-1 person. This is the **A-S theorem**.
- If $u\left(f\left(x_1, x_2\right), \ell\right)$: Start from any nonlinear tax and $t_2 \neq 0$, move to $t_2 = 0$ and adjust income tax: Pareto improving (see next extension).

$$t_2^* = -\frac{\gamma \widehat{v}_x^2 \left(\widehat{x}_2^2 - x_2^1\right)}{\lambda \left(n^1 \partial \widetilde{x}_2^1 / \partial q_2 + n^2 \partial \widetilde{x}_2^1 / \partial q_2\right)} \gtrless 0 \text{ as } \widehat{x}_2^2 \gtrless x_2^1$$

Numerator:
Alternatively, If x_2 is relatively complementary with leisure, the indirect tax on it should be positive and vice versa. Intuitively, if $\widehat{x}_2^2 > x_2^1$, then mimicking is made more difficult by imposing a tax on x_2.(.../..)

(.../..)
Suppose that $\widehat{x}_2^2 > x_2^1$. Starting at $t_2 = 0$, increase t_2 and adjust the income tax paid by the two types so that $dT^j = -x_2^j dt_2$. This will leave $v^j(.)$ as well as gvt revenue unchanged. However, $\widehat{v}^2(.)$ reduced, so incentive constraint relaxed: social welfare can be increased.

Atkison-Stiglitz, linear/nonlinear taxes on commodities

The Atkinson Stiglitz theorem also applies to non-linear commodity taxation.
Thus, under weak separability (and taste homogeneity) so that A-S theorem applies:

- Electricity should be taxed at the same rate as any other good
- Implementing a non-linear tax on electricity (e.g. a smaller tax rate for purchasers of smaller quantities) is not an efficient way to do redistribution

When weak separability does not hold so that A-S theorem fails, the optimal tax rates on commodities under an optimal non-linear income tax are given by:

- Atkinson Stiglitz (JPubE 1976) for **non-linear taxes on commodities!**
- Jacobs Boadway (JPubE 2014) for **linear taxes on commodities** which are, in the real world, more relevant.
- Christiansen (JPubE 1984) also **linear taxes on commodities** but (only) focuses on small welfare improving changes to commodity taxation (desirability of introducing small linear commodity taxes alongside optimal nonlinear income tax).

Atkison-Stiglitz revisited by Konishi (JPubE 1995), Laroque (EL 2005) and Kaplow (JPubE 2006)

Konishi (JPubE 1995), Laroque (EL 2005) and Kaplow (JPubE 2006) theorem:
With (any) nonlinear income taxes, if individuals have preferences that satisfy weak separability (and taste homogeneity), then it is always possible to generate a **Pareto improvement** by **eliminating differences in tax rates across commodities**.

This reform to uniform commodity taxes accompanied by a changes in the income tax system can be Pareto improving **without violating the revenue or incentive constraints**. That is, a given amount of redistribution can be achieved more efficiently.

The proof is based on Laroque (2005)

Setup (similar to before)

- n commodities indexed by i
- Each commodity i is associated with:
  - a production price p_i
  - a consumption price q_i
  - a tax rate t_i where $q_i = p_i(1 + t_i)$

Production prices are assumed to be exogenously fixed. This assumption requires a CRS technology together with a single non-produced input (efficiency units of labor).

Setup (cont'd)
(similar to before)

- Continuum of consumers, each indexed by j. Each consumer has:
  - productivity w^j (which is private information)
  - labor supply ℓ^j (which is also private information) et $y^j = w^j \ell^j$
  - consumption x_i^j of commodity i

- The consumer problem is:

$$\underset{\mathbf{x}^j,\ell^j}{Max}\ u\left(\mathbf{x}^j,\ell^j\right)$$

$$\text{s.to: } \sum_{i=1}^{n} p_i\left(1+t_i\right)x_i^j = y^j - T\left(y^j\right)$$

where $T(.)$ is the non-linear labor income tax schedule.

The proof is based on Laroque (2005)

Consider an initial policy $(T(.), t)$ where $t = (t_1, ..., t_n)$.
Let $v(y)$ denote the utility from consumption of a consumer with gross labor income y :

$$\begin{aligned} v(y) &= \underset{x_1,\ldots,x_n}{Max}\ u(x_1, ..., x_n) \\ \text{s.to} &: \sum_{i=1}^{n} p_i (1 + t_i) x_i = y - T(y) \end{aligned}$$

Define an alternative allocation of consumption
$\widetilde{x}(y) = (\widetilde{x}_1(y), ..., \widetilde{x}_n(y))$ by

$$\begin{aligned} \widetilde{x}(y) &= \arg \underset{x_1,\ldots,x_n}{Min} \sum_{i=1}^{n} p_i x_i \\ \text{s.to} &: v(x_1, ..., x_n) \geq v(y) \end{aligned}$$

Define a new non-linear income tax schedule $\widetilde{T}(.)$ by:

$$y - \widetilde{T}(y) = \sum_{i=1}^{n} p_i \widetilde{x}_i (y)$$

So $\widetilde{T}(.)$ is defined such that consumer with gross income y can afford $\widetilde{x}(y)$ in the absence of commodity taxes.

- Since consuming $\widetilde{x}(y)$ is the least costly way of getting utility $v(y)$ under the price vector p, then $v(y) = v(\widetilde{x}(y))$ is the maximal utility that can be reached with income $p \cdot \widetilde{x}(y) = y - \widetilde{T}(y)$.
- Thus, under $\left(\widetilde{T}(.), 0\right)$, a consumer with gross income y chooses $\widetilde{x}(y)$ and gets utility $v(y)$ from consumption.

Let $x(y)$ denotes the consumption of a worker with gross income y under the initial policy $(T(.), t)$.

By definition of $v(y)$, under $(T(.), t)$, a consumer with gross income y chooses $x(y)$ and gets utility $v(x(y)) = v(y)$ from consumption.

Thus, under both $(T(.), t)$ to $\left(\widetilde{T}(.), 0\right)$, a consumer who chooses to have gross income y^j gets utility $u^j\left(v\left(y^j\right), y^j / w^j\right)$.

- This implies that switching from $(T(.), t)$ to $\left(\widetilde{T}(.), 0\right)$ does not change the labor supply $\ell^j = y^j / w^j$ of any worker j.

We must have $y - \widetilde{T}(y) = \sum_{i=1}^{n} p_i \widetilde{x}_i(y) \leq \sum_{i=1}^{n} p_i x_i(y)$.

- Indeed, in the absence of commodity taxes, expenditures are minimized by $\widetilde{\mathbf{x}}(y)$ and not by $\mathbf{x}(y)$.
- The inequality is strict unless, initially, all commodities are taxed at the same rate (in which case $\mathbf{x}(y) = \widetilde{\mathbf{x}}(y)$).

Under the initial policy $(T(.), t)$, government revenue is equal to:

$$R = \int \left[T\left(y^j\right) + \sum_{i=1}^{n} p_i t_i x_i^j\left(y^j\right) \right] dj = \int \left[y^j - \sum_{i=1}^{n} p_i x_i^j\left(y^j\right) \right] dj$$

where the second equality was obtained by substituting consumer j's budget constraint for each j.

Under the new policy $\left(\widetilde{T}\left(.\right),0\right)$, government revenue is equal to:

$$\widetilde{R} = \int \widetilde{T}(y^j)dj$$

where, for each j, $y^j = w^j \ell^j$ is the same as under the old policy.
But, we know that for any j, $y^j - \widetilde{T}\left(y^j\right) \leq \sum_{i=1}^{n} p_i x_i^j\left(y^j\right)$.
Thus, if initially all goods are not taxed at the same rate (so that the inequality is strict), then $\int\left[y^j - \sum_{i=1}^{n} p_i x_i^j\left(y^j\right)\right]dj \leq \int \widetilde{T}\left(y^j\right)dj$ hence:

$$\widetilde{R} > R.$$

Switching from $(T(.), t)$ to $\left(\widetilde{T}\left(.\right),0\right)$ does not affect the utility of any individual, **but increases government revenue**.

Robustness and Intuition

The Atkinson Stiglitz theorem, as proved by Laroque (EL 2005), is very robust:

- It holds for any social welfare function.
- It does not require the Spence-Mirrlees condition to hold.
- It applies with either an intensive or an extensive margin to labor supply (or both).

Intuition:

- The government wants to redistribute from high to low productivity individuals.
- Under weak separability (and taste homogeneity), the consumption of commodities does not reveal any information about the productivity of an individual which has not already been revealed by his gross labor income.

Some **caution** regarding the policy relevance of Konishi (1995), Laroque (2005) and Kaplow (2006)'s result, see Boadway (JEL, 2010) and see also Boadway (2012, pp.64-65).

- If the Gvt could make the changes to the income tax system required to generate a Pareto-improving move to commodity tax unformity, there is seemingly little to preclude it from going to the fully optimal nonlinear income tax, in which case the Atkinson-Stiglitz theorem already applies.
- [Skip] As part of the proposed reform, the income tax can be changed at will by income-specific changes in tax liabilities. There is the potential for a Pareto improvement analogous to the potential Pareto improvement indicated by Kaldor-Hicks-Samuelson compensation test which meet several (serious) objections (see Boadway 2012, Chapter 2, pp.40-44) .

- From Konishi-Laroque-Kaplow approach, **we learn about the limits of the A-S theorem**. Failure of the theorem either because:
  - Non weakly separable preferences: relatively higher taxes on goods that are more complementary with leisure. However, limited evidence on that, no presumption that necessities should be taxed more lightly than luxuries.
  - No optimal income tax in place: the direction of deviation from uniformity is not obvious, it depends on the form of the income tax in place and assumptions about the prospects of income tax reform.
  - ⇒ One can argue that uniformity is then a reasonable policy choice. "The costs of deviating from the optimum are likely strictly convex, so if one does not know in which direction to deviate, expected losses are minimized by implementing uniformity." (Boadway, 2012, p.65)

Besides the requirement for optimal income tax to be in place, there are several other **caveats or extensions** to the A-S theorem:

- different needs for particular commodities across the population, differences in tastes or preferences
- allocation of time by individuals to nonmarket labor or to consumption of purchased goods [Skip, see Boadway (2012, pp.69-71)]
- emission of an externality by one or more goods
- imperfect substituability of different skills of labor
- wage uncertainty
- dynamic optimal taxation

In each case, assume weak separability of goods from leisure in order to focus on other possible violations of the A-S theorem.

Extension: Needs or Endowments

Assume separable utility $u\left(f\left(x_1, x_2 - r\right), \ell\right)$ where r is either
1. need for good x_2 (Rowe and Woolley CJE, 1999; Boadway and Pestieau, 2003)
2. initial endowment of x_2 (Cremer, Pestieau and Rochet, IER 2001), or
3. taste factor (Saez JPubE 2002, Blomquist and Christiansen FinanzArchiv, 2008)

Let $r = \{r^1, r^2\}$, high and low need.
Four individual types $\left\{w^i, r^j\right\}$, $i, j = 1, 2$
Two cases:

1. w not observable, r observable
2. w and r not observable

Case I: w not observable to the Gvt and r observable

Assume r and w imperfectly correlated
Population can be divided into two identifiable need groups:
$\left\{w^1, r^1; w^2, r^1\right\}$ $\left\{w^1, r^2; w^2, r^2\right\}$
Needs serve as a signal/**tag** (Akerlof AER,1978)

- Incentive constraint only relevant within each needs group (e.g., tagging)
- Lump-sum redistribution between needs groups

Let $i = 1, 2$ index wages and $j = 1, 2$ index needs

$$\Rightarrow t_2 \sum_i \sum_j n^{ij} \frac{\partial \widetilde{x}_2^{ij}(q_2)}{\partial q_2} = \sum_j \frac{\gamma_j}{\lambda} \widehat{v}_f^{2j} \widehat{f}_c^{2j} \left(x_2^{1j} - \widehat{x}_2^{2j} \right) = 0$$

since within needs groups $x_2^{1j} = \widehat{x}_2^{2j}$ $\Rightarrow$ **A-S theorem applies**.

- Separate nonlinear income tax within each need group.

Case I: w not observable and r observable (cont'd)

Remark: Lump-sum redistribution between needs groups; the lump-sum transfer between the two groups should be sufficient

1. to compensate for differences in needs per capita between the two groups,
2. to compensated for differences in the share of high-skilled workers between the two groups (transfer from group with a larger proportion of high- to low-skilled persons to the group with the smaller proportion).

Case II: w, r not observable

If needs not observable: **A-S generally does not apply**.

Assume w, r perfectly correlated: e.g. two types (w^1, r^1), (w^2, r^2)
If needs are increasing with wage rate: differential commodity tax on good x_2 alongside an optimal nonlinear income tax.

Intuition: Higher wage persons will demand more of good x_2 for a given level of income, so taxing x_2, relaxes the incentive constraint. The opposite applies if needs fall with wage rate.

Case II: w, r not observable (cont'd)

A-S theorem will still **not apply** and more complicate to characterize the optimal commodity tax system if:

- needs apply to more than one good
- needs are only imperfectly correlated with wage rate

Heterogeneous needs/preferences for leisure

Needs can apply to leisure rather than goods; $u(f(x_1, x_2), x_0 - r)$ where r is (heterogeneous) need for leisure

- A-S theorem applies, regardless of whether r is observable
- Because of separability: the demand for commodities depends only on disposable income: a high-wage person mimicking the income of a low-wage person consume the same bundle, so differential commodity taxes serve no purpose.
- Also true for more general form of interraction between x_0 and r, e.g. rx_0.
- Optimal nonlinear income tax affected by leisure needs:

1. If needs observable: they serve as a tag and $\neq$ tax schedule in $\neq$ needs groups.
2. If needs not observable: two-dimensional screening problem with ambiguous pattern of ICC.

Heterogeneous preferences/tastes

Those preferences are related to differences in needs.

Results:

- Higher taxes on goods demanded more by higher-skilled persons (Saez 2002, Blomquist and Christiansen 2008).

Intuition (as earlier with nonseparable preferences): such a tax weakens the incentive constraints and facilitate greater redistribution to lower skilled persons.

Heterogeneous preferences/tastes

Those preferences are related to differences in needs.

Results:

- Intertemporal case (goods consumed in $\neq$ periods are $\neq$ goods; heterogeneous preferences = heterogeneous utility discount rates): higher tax on future consumption if high-skilled have lower utility discount rates (Saez 2002), $\sim$ tax on capital income (savings)

Intuitively, if utility discount rate lower for higher skilled persons, they will have higher saving rates. $\Rightarrow$ **tax on capital income** serves as an indirect way of redistributing from the high-skilled to the low-skilled. [See topic “Capital taxation"]

Environmental Externality: Pigouvian Taxation/ Externalities

Suppose x_2 is now a dirty good and weakly separable utility $u\left(f\left(x_1, ..., x_n\right), x_0, e\left(X_1\right)\right)$ where $e\left(X_2\right)$ is the quality of the environment, decreasing in aggregate consumption of dirty good X_2

If m is the population indexed by j, then $X_2 = \sum_{j=1}^{p} x_1^j$.

For illustration, I simplify:

$$\text{Utility: } h(x_1) + k\left(x_2\right) - \frac{y}{w} + e$$

where $e = \overline{e} - \delta(n^1 x_1^1 + n^2 x_1^2)$ (externality)

Nonlinear tax on y, excise tax t_2 on x_2 as before.

Individual choice of x_1, x_2, given c, y:

$$\underset{x_2}{Max}\ h(c - q_2 x_2) + k(x_2) - y/w + e \Rightarrow x_2(q_2, c)$$

$$v(q_2, c, y) + e,\ v_{q_2} = -x_2 h',\ v_{q_2} = h',\ v_y = 1/w$$

For mimicker: $\widehat{x}_2^2\left(q_2, c^1\right)$, $\widehat{v}_2\left(q_2, c^1, y^1\right) + e$
Note: Given separability, $\widehat{x}_2^2 = x_2^1$.

First-Best Government Policy

In a first-best world, the optimal corrective/Pigouvian tax t_2 is equal to $\sum$ of marginal damages to all individuals evaluated at their own marginal utilities of income.

$$\underset{c^j, y^j, t_2}{Max}\, \rho^1 n^1 v^1(q_2, c^1, y^1) + \rho^2 n^2 v^2(q_2, c^2, y^2) + \overline{n}e$$

s.to.

$$n^1 \left(y^1 - c^1 + t_2 x_2^1 \left(q_2, c^1\right)\right)$$
$$+n^2 \left(y^2 - c^2 + t_2 x_2^2 \left(q_2, c^2\right)\right) = R \quad \text{(mult. } \lambda^{FB}\text{)}$$

where $\overline{n} = \rho^1 n^1 + \rho^2 n^2$

First-Best Government Policy (cont'd)

FOCs yield:

$$\lambda^{FB} = \rho^1 v_y^1 = \rho^2 v_y^2$$
$$t_2^{FB} = \frac{n^1 \delta}{v_y^1} + \frac{n^2 \delta}{v_y^2}$$

- Equality of marginal social utility of incomes. The Gvt imposes redistributive lump-sum taxes to equalize the marginal social utilities of income among all individuals.
- Pigouvian tax $t_2^{FB} = \sum$ marginal damages evaluated by individuals **at their own marginal utilities of income** (no social welfare weights).

Second-Best Government Policy

The Gvt observes only individual incomes and market transactions.

- Cremer, Gahvari, Ladoux (*JPubE* 1998): Given weakly separable utility function (as above), no differential taxes should be applied to nonpolluting goods, let us denote the common tax rate t.
- Boadway and Tremblay (Asia-Pac.JAE 2008): tax on dirty good x_2 plays a dual role, partly as Pigouvian corrective tax (t_2^c) and partly as the component of the commodity tax system used for revenue-raising purposes (t_2^R). The Pigouvian tax $t_2^c = \sum$ marginal **social** damages evaluated at the marginal utility of Gvt revenue, where the latter is an average of marginal social utilities of income.

Second-Best Government Policy (cont'd)

- Kaplow (IER 2012): A Pigouvian tax on the polluting good. With a non-linear tax on labor income and weakly separable preferences in labor (and taste homogeneity), it is possible, adjusting income tax, to move from any set of commodity taxes to first-best Pigouvian taxes (wherein each tax equals marginal harm) in a manner that generates a Pareto improvement.

Second-Best Government Policy (cont'd)

$$\underset{c^j,y^j,t_2}{Max}\, \rho^1 n^1 v^1(q_2, c^1, y^1) + \rho^2 n^2 v^2(q_2, c^2, y^2) + \overline{n}e$$

$$\text{s.to. } v^2\left(q_2, c^2, y^2\right) \geq \widehat{v}^2\left(q_2, c^1, y^1\right)$$

$$n^1\left(y^1 - c^1 + t_2 x_2^1\left(q_2, c^1\right)\right)$$

$$+n^2\left(y^2 - c^2 + t_2 x_2^2\left(q_2, c^2\right)\right) = R \quad (\text{mult. } \lambda^{SB})$$

$$\text{where } \overline{n} = \rho^1 n^1 + \rho^2 n^2$$

Second-Best Government Policy (cont'd)

From the FOCs, we obtain, using $\widehat{x}_2^2 = x_2^1$:

$$t_2^c = \frac{\overline{n}\delta}{\lambda} = \frac{\left(\rho^1 n^1 + \rho^2 n^2\right)\delta}{\lambda^{SB}}$$

- Pigouvian tax equals sum of marginal social damages (using social weights ρ^1, ρ^2) in terms of government revenue λ where

$$\lambda^{SB} = \frac{n^1 \rho^1 v_y^1 + n^2 \rho^2 v_y^2}{n^1 + n^2}$$

Interpretation of Pigouvian tax

Rewrite Pigouvian tax as:

$$t_2^c = \frac{n^1 \delta}{\lambda^{SB}/\rho^1} + \frac{n^2 \delta}{\lambda^{SB}/\rho^2}$$

Since $\rho^1 v_y^1 > \rho^2 v_y^2$ (marginal social utilities of income),

$$\frac{\lambda^{SB}}{\rho_1} < v_y^1, \ \frac{\lambda^{SB}}{\rho_2} > v_y^2$$

$$\lambda^{SB} = \frac{n^1 \rho^1 v_y^1 + n^2 \rho^2 v_y^2}{n^1 + n^2}$$

- Pigouvian SB tax puts more weight on marginal social damages to low-wage persons than high-wage persons. Reminiscent of Sandmo (JPubE 2006) for global externalities in absence of optimal international transfers, Boadway and Tremblay (Asia-Pac.JAE 2008).
- Pigouvian SB tax plays some redistributive role, Boadway and Tremblay (Asia-Pac.JAE 2008)

Non weakly separable preferences in the n goods$\Rightarrow$

- A-S theorem does not apply, no uniform optimal taxes on nonpolluting goods; higher taxes on goods more substituable with leisure.
- The optimal tax on the polluting good x_2 includes a corrective Pigouvian component and a component reflecing the relative complementary of x_2 with leisure compared with the $n-1$ other goods.

Skills are imperfect substitutes

- Stiglitz (JPubE 1982) result for two-skill-type optimal nonlinear income tax model: if high- and low-skilled workers are imperfect substitutes, optimal marginal tax rate on high-skilled persons should be negative. Intuitively, a negative T' increases labor supply by the high-skilled $\Rightarrow$ increases in the low-skilled wage rate in general equilibrium as the ratio of high- to low-skilled labor rises.
- Naito (JPubE 1999): in two-good extension (i.e. adds one good to Stiglitz (1982)), higher tax should apply to more skill-intensive good (A-S fails).
- Saez (JPubE 2004): If workers allowed to change occupations, A-S theorem applies even with endogenous relative wages, and zero tax rate at the top restored.

Naito (JPubE 1999): Setup

- We assumed that production prices are exogenously fixed. This requires constant returns to scale together with a single non-produced input (efficiency units of labor).
- Gauthier Laroque (JPubE 2009) showed that the A-S theorem also holds under decreasing returns to scale with a single non-produced input provided that profits can be taxed away.
- Naito (1999) showed that, with several non-produced inputs, A-S theorem no longer holds.

Assumptions of Naito (1999):

- High-skilled and low-skilled labor are not perfect substitutes.
- They should be seen as two different non-produced inputs.
- There are two goods produced under constant returns to scale from high-skilled and low-skilled labor.
- The production of one good is more skill intensive than that of the other.

Naito (JPubE 1999): Result

Theorem: Under weak separability, taste homogeneity and with a non-linear income tax used for redistribution, the optimal policy is to tax the high-skilled-labor intensive commodity at a higher rate than the low-skilled-labor intensive commodity.

Intuition:

- The tax raises the demand for the low-skilled-labor intensive commodity.
- This raises the wage of low-skilled workers relative to that of high-skilled workers.
- This relaxes the incentive compatibility constraint of the high-skilled.
- Indeed, with a smaller wage differential, a high-skilled worker is less tempted to mimic a low-skilled worker (as this requires working harder than with a larger wage differential).

Saez (JPubE 2004)

Saez (2004) restores AS theorem with endogenous wages when workers can choose intensive (effort) margin and occupations.

- High-skilled workers can choose to work in low-skilled jobs and earn the same incomes as the latter.
- An indirect tax aimed at increasing low-skilled wages $\Rightarrow$ high-skilled workers to switch occupations.

$\Rightarrow$ No advantage anymore to using commodity taxes in addition to a nonlinear income tax.

Wage uncertainty (Cremer and Gahvari, EJ 1995)

Cremer and Gahvari (EJ 1995): study validity of A-S theorem when **wage rates are uncertain, individuals make some consumption decisions** (consumer durables, e.g. housing) **before they know their wage rate.**

- Dynamic setting not needed: intuition of their result can be shown in static model in which decisions are taken sequentially
- Nonlinear tax on ex post incomes
- commodity taxes on consumer purchases when they are made

Wage uncertainty (cont'd)

- If all goods' purchases made ex post (i.e. no durables) $\Rightarrow$ A-S theorem applies; no differential commodity taxes if goods weakly separable from leisure in the utility function.
- If some goods purchased before uncertainty resolved $\Rightarrow$ A-S theorem fails. Goods purchased ex ante bear a lower tax rate.

Intuitively, inducing **all** persons to increase their consumption of durable good makes it more difficult for those who turn out to have higher skills ex post to mimic those with low skills, since ex post consumption requirements are higher.

Wage uncertainty (cont'd)

- Preferential tax treatment of durables to offset the excessive precautionary saving to self-insure against wage uncertainty.

Cremer and Gahvari argue : this gives some justification for preferential tax treatment of housing and consumer durables; the imputed rent to the consumer goes untaxed.

Wage uncertainty (cont'd)

- Cremer an Gahvari's argument **closely related to those in favor of taxing capital income in a dynamic setting with wage uncertainty** (Golosov, Tsyvnski and Werning 2007; Diamond 2007).

Dynamic optimal nonlinear taxation

"New dynamic public finance": extension of Mirrlees model to intertemporal setting where

- heterogeneous individuals live for several (maybe infinite) periods,
- consume and supply labor at each period,
- carry saving from one period to the other.

Dynamic analogue of A-S theorem in Golosov, Kocherlakota and Tsyvinski (ReStud 2003).

Dynamic optimal nonlinear taxation (cont'd)

Dynamic analogue of A-S theorem in Golosov, Kocherlakota and Tsyvinski (2003):

- lifetime utility is discounted sum of per period utilities
- utilities are weakly separable function of labor and the vector of per period consumption
- skill evolves arbitrarily with each period, and a person's own skill level is revealed privately at the beginning of the period
- Gvt max. weighted sum of individual lifetime utilities with (arbitrary) weights based on initial skills

Dynamic optimal nonlinear taxation (cont'd)

- Gvt can **fully commit** ex ante to an optimal tax system
- Gvt per period resource constraint
- The gvt can condition policy in each period **on the full history of information learned in all previous periods** (including past skill levels)
- set of incentive constraints. **Individual incentive constraints apply on a lifetime basis**: each person must weakly prefer her lifetime consumption-income allocation to that of any other
- Problem solved by direct mechanism by which the Gvt chooses consumption bundles and income for all individuals in each period.

Dynamic optimal nonlinear taxation (cont'd)

Main results (Golosov, Kocherlakota and Tsyvinski ReStud, 2003):

- Characterization of intertemporal wedges i.e. the extent and direction of distortions of intertemporal consumption/saving decisions that influence the tax on savings or capital income
- Taxation of goods in each period: Per period analogue of A-S theorem applies. In each period, the MRS between any pair of goods equals its MRT for all persons. Thus **no indirect taxation**.

Mixed Taxation: Policy Recommendation from the Mirrlees Review

In 2010 and 2011, a large group of top-notch academics made a detailed assessment of the characteristics of a good tax system for an open developed economy, with a focus on the UK.
The Mirrlees review recommends:

- Uniform VAT rates across all commodities;
- Except for child care which should benefit from a much lower rate.

Outline

Akerlof, G.A. (1978), The economics of "tagging" as applied to optimal income tax, welfare programs, and manpower training, American Economic Review, 68, 8–19.

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

Atkinson, A.B. and Stiglitz, J.E. (1980), Welfare implications of the taxation of savings, Economic Journal, 90, 529–549.

Blomquist, S. and Christiansen, V. (2008), Taxation and heterogeneous preferences, FinanzArchiv, 64, 218–244.

Boadway, R. (2010), Efficiency and redistribution: An valuative review of Louis Kaplow's The Theory of Taxation and Public Economics, Journal of Economic Literature, 48, 964–979.

**Boadway, R. (2012), From Optimal Tax Theory to Tax Policy, Retrospective and Prospective Views, CESifo Munich lectures in economics, MIT Press, Chapter 3.

Boadway, R. and Pestieau, P. (2003), Indirect taxation and redistribution: The scope of the Atkinson-Stiglitz theorem. In Richard Arnott, Bruce Greenwald, Rabi Kanbur and Barry Nalebuffs, eds., Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz, Cambride: MIT Press, 387–403.

Boadway, R. and J-F. Tremblay (2008), Pigouvian taxation in a Ramsey world, Asia-Pacific Journal of Accounting and Economics, 15, 183–204.

Christiansen, V. (1981), Evaluation of public projects under optimal taxation, Review of Economic Studies, 48(3), 447–457.

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

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

Cremer, H., Gahvari, F. (1995), Uncertainty, optimal taxation and the direct versus indirect tax controversy, Economic Journal, 105, 1165–1179.

Cremer, H., Gahvari, F. and Ladoux, N. (1998), Externalities and optimal taxation, Journal of Public Economics, 70, 343–364.

Cremer, H., Pestieau, P. and J-C. Rochet (2001), Direct versus indirect taxation: The design of the tax structure revisited, International Economic Review, 42, 781–799.

Diamond, P. (2007), Comment on Golosov et al., NBER Macroeconomic Annual, 2006: 365–379.

Edwards, J., Keen, M. and M. Tuomala (1994), Income tax, commodity taxes and public good provision: A brief guide, FinanzArchiv, 51, 472–497.

Gauthier, S. and Laroque, G. (2009), Separability and Public Finance, Journal of Public Economics, 93, 1168–1174.

Golosov, M., Kocherlakota, N. and Tsyvinski, A. (2003), Optimal indirect capital taxation, Review of Economic Studies, 70, 569–597.

Golosov, M. Tsyvinski, A. and Werning, I. (2007), New dynamic public finance: A user's guide, NBER, Macroeconomics Annual, 2006, 317–363.
Harberger, A.C. (1964), Taxation, resource allocation and welfare. In John Due, ed., The Role of Direct and Indirect Taxes in the Federal Revenue System, Princeton: Princeton University Press, 25–80.
**Jacobs, B. and Boadway, R. (2014), Optimal linear commodity taxation under optimal non-linear income taxation, Journal of Public Economics, 117, 201–210.
Kaplow, L. (2006), On the undesirability of commodity taxation even when income taxation is not optimal, Journal of Public Economics, 90, 1235–1250.
Kaplow, L. (2012), Optimal control of externalities in the presence of income taxation, International Economic Review, 53(2), 487–509.

Konishi, H. (1995), A Pareto-improving commodity tax reform under a smooth non-linear income tax, Journal of Public Economics, 56, 413–446.

**Laroque, G. (2005), Indirect taxation is superfluous under separability and taste homogeneity: A simple proof, Economics Letters, 87, 141–144.

Mirrlees, J. (1976), Optimal Tax Theory: A Synthesis, Journal of Public Economics, 6, 327–358.

Nava, M., Schroyen, F. and M. Marchand (1996), Optimal fiscal and public expenditure policy in two-class economy, Journal of Public Economics, 61, 119–137.

Naito, H. (1999), Re-examination of uniform commodity taxes under a non-linear income tax system and its implications for production Efficiency, Journal of Public Economics, 71, 165–188.

Rowe, N. and Woolley, F. (1999), The efficiency case for universality, Canadian Journal of Economics, Revue Canadienne d'Economique, 32, 613–629.
Saez, E. (2002), The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes, Journal of Public Economics, 83, 217–230.
Saez, E. (2004), Direct or indirect tax instruments for redistribution: Short-run versus long-run, Journal of Public Economics, 88, 503–518.
Sandmo, A. (1976), Optimal taxation: An introduction to the literature, Journal of Public Economics, 6, 37–54.

Sandmo, A. (2006), Global public economics: Public goods and externalities, Public Economics/Economie Publique, Revue de l'Institut d'Economie Publique,18–19 (http://economiepublique.revues.org/4282?file=1)

Stiglitz, J.E. (1982), Self-selection and Pareto efficient taxation, Journal of Public Economics, 17, 213–240.