# Optimal tax and transfer policy
## (non-linear income taxes and redistribution)

March 2, 2016

# Non-linear taxation

- So far we have considered linear taxes on consumption, labour income and capital income (briefly). With a linear tax, taxes paid divided by the tax base is constant as the base varies.
- Direct taxes are often non-linear with transfers to those who do not participate in the labour market, or who have low earnings in the labour market. The marginal tax rate varies with income earned for those who participate, usually it is increasing in earnings.
- A non-linear tax is used to redistribute income across households. The optimal non-linear tax must balance distributional and efficiency concerns.
- The efficiency loss associated with non-linear taxes - (excess burden) depends on two behavioral responses; (i) whether or not individuals wants to work (participation, the extensive margin) and (ii) their work effort if they decide to participate (intensive margin)

# A taxonomy of taxes and transfers

- $y$ is pre-tax income and $T(y)$ is the tax paid if income is $y \Longrightarrow$ consumers after tax income is given by $(y - T(y))$
- $T'(y)$ is the marginal tax rate associated with income $y$: An individual keeps $1 - T'(y)$ of earning one extra NOK in income.
- The marginal tax rate matters for the choice of work effort (hours worked (training education)) for those who are active.
- A tax scheme is progressive if the tax rate increases as the tax base (income) increases: if the marginal tax rate is higher than the average tax rate.
- Some measures of progressively
  - *tlp*: The elasticity of the tax bill with respect to pre-tax income: $\frac{T'(y)}{T(y)/y}$
  - *rip*:The elasticity of the residual income with respect to pre-tax income: $\frac{1-T'(y)}{1-T(y)/y}$

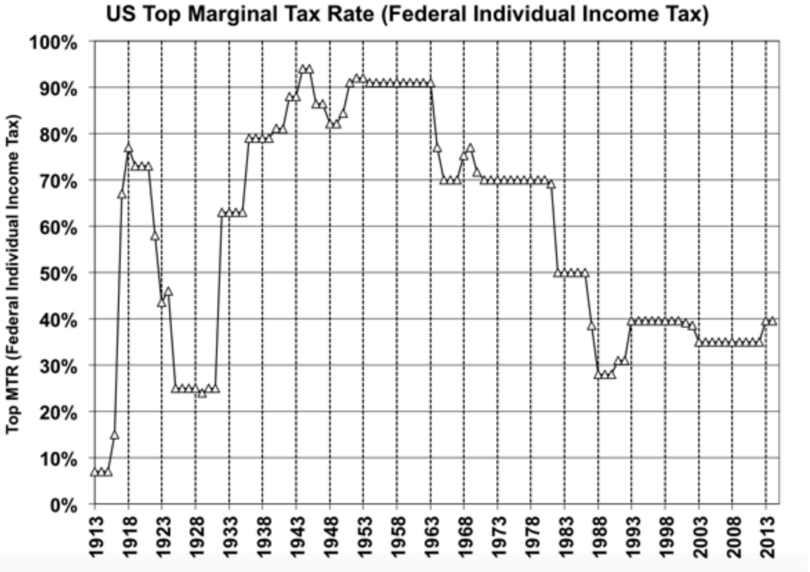
# A taxonomy of taxes and transfers

- The participation tax rate is defined by $\tau^p(y) = \frac{(T(y)-T(0))}{y}$, an individual that goes from inactivity to income $y$ keeps a fraction $1-\tau^p(y)$ of that increase in income: $y - T(y) - T(0) = (1-\tau^p(y))\, y$. (if $y = 300$, $T(300) = 100$ and $T(0) = 100$ we have $\tau^p(300) = \frac{2}{3}$)
- The participation tax rate matters for the incentives to participate in the labour market.
- We can express net earnings in terms of the participation tax rate

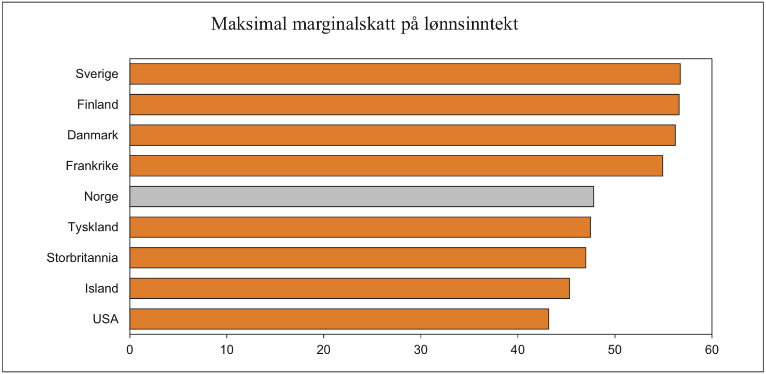
$$y - T(y) = -T(0) + y - (T(y) - T(0)) = -T(0) + y(1-\tau^p(y))$$

- Break even income, $(y^*)$: $T(y^*) = 0$

# Marginal tax rates



# Marginal tax rates

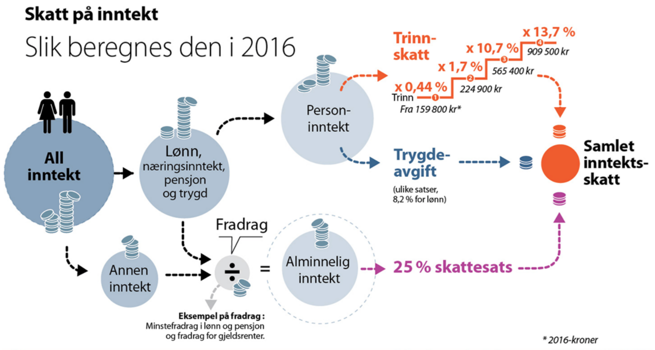



Figur 2.4 Høyeste marginalskatt på lønnsinntekt. Noen utvalgte land i 2013. Prosent

Kilde: OECD Tax database.

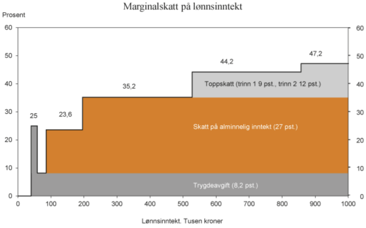
# Marginal tax rates

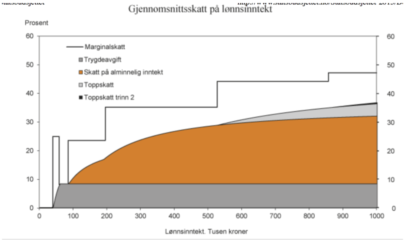
There was a “tax reform” in Norway this year. The tax rate on “ordinary income” was lowered from 27 to 25 %. More steps in the surtax.



# Marginal tax rates

2015



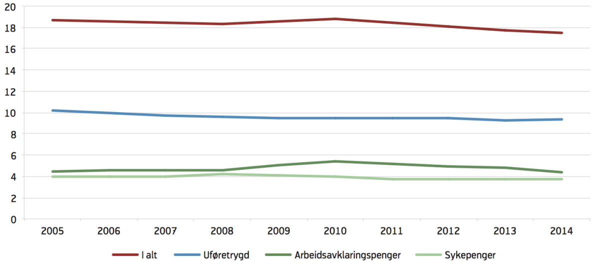
# Participation tax rates

- The literature on optimal income taxation (Mirrlees) has focused on the intensive margin; how marginal tax rates reduce the incentives to work (train and educate).
- Empirical studies find low labour supply elasticity for those who work.
- Higher elasticity on the extensive margin: the “not-participate” “participate” choice is more sensitive to changes in transfers and taxation.
- Not only the “participation tax rate” that matters, but also the requirements, conditions, that are associated with receiving welfare benefits (activation)

# Participation

Close to 25 % of individuals in working age (16 - 74), around 5 % receive social benefits (means tested economic support) and around 20 % health related compensation for lost work capacity (100% or less)

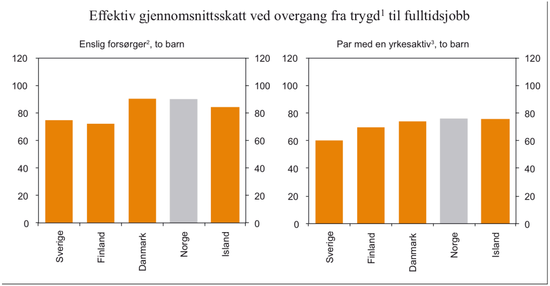
**Figur 1.** **Andel av befolkningen 18–66 år som mottar sykepenger, arbeidsavklaringspenger eller uføretrygd ved utgangen av året, justert for dobbelttellinger.* Prosent**




*Kilde:* NAV

* foreløpige tall for 2014

# Participation tax rates




(2) Comparing 0 income with earning 67% of average income
(3) Comparing 0 income with earning 100% of average income

# Optimal income taxes with fixed income (no behavioral response)

- All individuals have the same strictly increasing and concave utility function $u(c)$.
- Income $y$ is fixed (exogenous) and consumption is equal to income after tax: $c = y - T(y)$.
- Government maximizes Utilitarian objective: $\int_0^{\bar{y}} u(y - T(y))h(y)dy$, where $h(y)$ is the distribution (pdf) of income over the interval of income in the economy $[0, \bar{y}]$.
- Budget constraint $\int_0^{\bar{y}} T(y)h(y)dy = R$ (multiplier $\lambda$)
- Lagrangian $L = \int_0^{\bar{y}} (u(y - T(y)) + \lambda [T(y) - R])\, h(y)dy$

# Optimal income taxes with fixed income (no behavioral response)

Lagrangian $L = \int_0^{\bar{y}} (u(y - T(y)) + \lambda [T(y) - R])\, h(y) dy$

- F.o.c:

$$
\begin{aligned}
0 &= \frac{\partial L}{\partial T(y)} = -u'(y - T(y)) + \lambda h(y) \\
0 &= u'(y - T(y)) = \lambda
\end{aligned}
$$

$$\Longrightarrow y - T(y) = c$$

$$c = \bar{y} - R$$

- Equalization of after tax income $\Longrightarrow$ 100% marginal tax rate

# Mirrlees (1971)

## Optimal taxation with endogenous labour supply and heterogeneous income potential

- individuals maximize $u(c, L)$ s.t. $c = wL - T$, $w =$ wage rate, $L$ is labour supply and $T$ are taxes.
  - individuals differ in wages (abilities) which is distributed with density $f(w)$.
  - Government maximizes a social welfare function $W(u(c, L))$ (increasing and concave)

$$SWF = \int W\left(u(c, L)f(w)\right) dw$$

Subject to a budget constraint

$$\int T(wL)f(w)dw = R$$

and a behavior constraint (IC)

$$w(1 - T')u_c + u_L = 0$$

# Mirrlees1971

## Main results

- Mathematically a complex problem (choose an optimal tax function)
- Relatively few general insights (unless we put more structure on parameters and functions)
- Concave *SWF* implies $T < 0$ for individuals with low wages, $T > 0$ for higher individuals with higher income; the degree of redistribution depends on the concavity of $W$ and the elasticity of labour supply (for those with high earning capacity)
- Never optimal to have $T' < 0$: tax liability should increase in wage.
- $T' = 0$ for the individual with *highest* ability ($w$) (marginal tax rate should be zero, not the average tax rate)

# Two type version of Mirrlees

- The main incentive problem in the Mirrlees framework is that a tax on income induce individuals to reduce their income (work less) - “they mimic individuals with lower income potential (ability).
- The policy question is to find the optimal tax scheme taking into account this incentive problem (a mimicking constraint).
- Possible to illustrate the problem in a two-type model. Two papers on the syllabus that covers this - Stiglitz (New-new) and Boadway and Keen).

# Two types

- The households preferences over consumption and labour (leisure) is given by $U(C, L)$, with $U_C > 0$ (the marginal utility of consumption is positive) and $U_L < 0$.
- $C_i = Y_i = w_i L_i$ with $w_H > w_L$ ($H =$ high income potential household)
- The household problem max $U(C, L)$ s.t. $wL = C$ :
- MRS $=$ MRT $= \frac{dC}{dL}_{u=U} = \frac{-U_L}{U_C} = w = \frac{dC}{dL}_{in\ \ market}$
- Rewrite utility in terms of what is observable for the government $Y$ and $C$. Since $L = \frac{Y}{w}$ we obtain $U(C, L) \equiv U(C, \frac{Y}{w}) = V(C, Y)$.
- Key observation: Indifference curves increases in the $Y, C$ space, and the indifference curve gets flatter the higher the wage is: $\frac{-V_Y}{V_C} = \frac{-U_L}{wU_C}$.
- Hence without any taxes optimality requires that $\frac{-V_Y}{V_C} = \frac{-U_L}{wU_C} = 1$ (MRS=MRT)

# Two types

## government problem

- The government imposes taxes contingent on observed income $T(Y)$.
- The government has a budget constraint: $T(Y_L) + T(Y_H) \geq R$, which can be written as $Y_H - C_H + Y_L - C_L \geq R$.
- Suppose the government implements consumption and income directly to each type of household: $\{C_L, Y_L\}$ to the household if it says it is of the $L-$type and the bundle $\{C_H, Y_H\}$ if it is of the $H-$type. This is called a direct mechanism in contract theory.
- Note that if an $L-$type takes the bundle designed for her she has to work $\frac{Y_L}{w_L}$ hours while a $H-$type has to work only $\frac{Y_L}{w_H}$ hours: So it is not possible to implement equal consumption here; $C_L$ must be lower than $C_H$ otherwise a $H-$type will pretend to be $L$ and consume a lot of leisure.

# Two types

- To characterize Pareto-optimal tax structure we maximize one types utility given the following constraints:
  - The government must cover its budget
  - Each type must prefer the bundle government offers that type
  - The utility of the household that we are not maximizing the utility of cannot drop below a certain level.

# Two types

Max $V^H(C_H, Y_H)$ subject to

1. $V^L(C_L, Y_L) \geq v$
2. $Y_H - C_H + Y_L - C_L \geq R$
3. $V^H(C_H, Y_H) \geq V^H(C_L, Y_L) = V^{H(L)} =$ utility of $H$if she takes the bundle for $L$
4. $V^L(C_L, Y_L) \geq V^L(C_H, Y_H) = V^{L(H)}$

- The two last equations are the incentive constraints, the self selection constraint in the information economics jargon. In this simple two type model the constraint will always bind for one of the types and not for the other, so there is a separating equilibrium - they choose different boundles. This is not always true in a more general model.

# Two Types

- The first order condition for the Lagrangian ($G$) of this problem with multipliers $(\mu, \gamma, \lambda_H, \lambda_L)$ are given by

$$\frac{\partial G}{\partial C_L} = \mu V^L_{C_L} - \lambda_H V^{H(L)}_{C_L} + \lambda_L V^L_{C_L} - \gamma = 0 \quad (1)$$

$$\frac{\partial G}{\partial Y_L} = \mu V^L_{Y_L} - \lambda_H V^{H(L)}_{Y_L} + \lambda_L V^L_{Y_L} + \gamma = 0 \quad (2)$$

$$\frac{\partial G}{\partial C_H} = V^H_{C_H} + \lambda_H V^H_{C_H} - \lambda_L V^{L(H)}_{C_H} - \gamma = 0 \quad (3)$$

$$\frac{\partial G}{\partial Y_H} = V^H_{Y_H} + \lambda_H V^H_{Y_H} - \lambda_L V^{L(H)}_{Y_H} + \gamma = 0 \quad (4)$$

- The most natural interesting case is where $\lambda_H > 0$ and $\lambda_H = 0$. The utility of $L$ is set so high that Manipulating these constraints and we find that (derived in my note and in Stiglitz)

$\frac{-V_{Y_H}}{V_{C_H}} = \frac{-U_L}{w_H U_C} = 1$ and $\frac{-V_{Y_L}}{V_{C_L}} = \frac{-U_L}{w_L U_C} < 1$.

- In terms of taxation (not direct implementation) this means that $L$ must face a positive marginal tax rate, and the $H-$type a 0 marginal tax rate. No marginal tax rate at the top, but a high average tax rate.

# Extensive and intensive margin: Saez 2002.

- The Mirrlees model focus on the intensive distortion of taxation; all individuals work and the issue is that a labour income tax distorts how much they work.
- Empirics show that the extensive margin is more important; an income tax (transfer) affects the decision to work or not.
- Very important paper by Saez considers ho an income tax affects both the extensive and intensive (Mirrlees) margin.
- The main result is that if the extensive margin is important for low income earners it may be optimal to have a negative marginal tax rate for lower incomes (EITC).

# Extensive and intensive margin: Saez 2002.

**Extensive margin**

- Consider a discrete model with $I+1$ occupations with wages $w_0 = 0 < w_1 < w_2 < ..$ and $w_0$ is the wage earned when not working.
- Tax/transfer $T_i$ when earning $w_i$ , $c_i = w_i - T_i$
- Only participation choice- individuals compare $c_i$ and $c_0$: $c_i - c_0 = (1 - \tau_i^p)w_i$ (by assumption not possible to move between occupations, no intensive margin).
- A fraction $h_i$ earns $w_i$, with $\sum h = 1$.
- Extensive responsiveness: % change in occupation $i$ for 1% increase in after tax consumption $i$ relative to no work: $\eta_i = \partial h_i / \partial (c_i - c_0) \cdot (c_i - c_0)/h_i$
- Social Welfare summarized by social marginal welfare weights at each earnings level $g_i$: transferring one dollar to someone with wage $w_i$ is worth
- These welfare weights sum to 1 and it is reasonable to assume they decline with consumption level: $g_0 > g_1 > g_2 > ..g_a = 1 > g_{a+1} > g_{a+2}....$
- Welfare weights will typically depend on the tax system, they are endogenous. If the tax system is such that $c_i = c_0$ for all $i$ would also be reasonable to have $g_i = 1$ for all $i$.

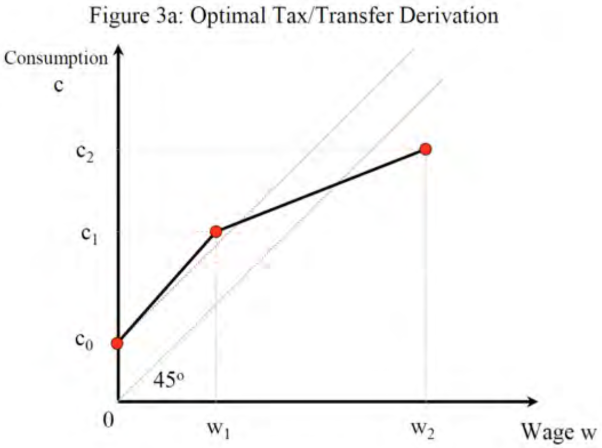
# Extensive and intensive margin: Saez 2002.

- Optimal tax rates (only extensive margin)

$$\frac{\tau_i^p}{1-\tau_i^p} = \frac{T_i - T_0}{c_i - c_0} = \frac{1}{\eta_i}(1-g_i)$$

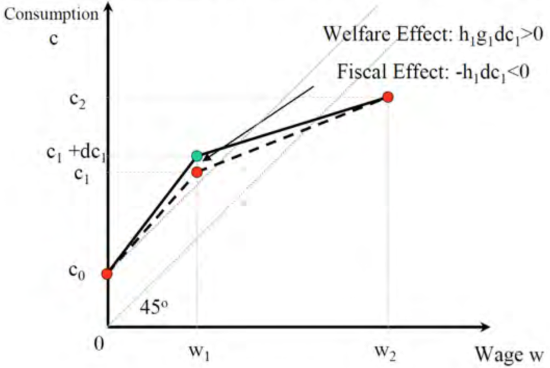
- If $g_i > 1$ we have $T_i - T_0 < 0 \Longrightarrow$ higher positive transfer to wage earners than those without work: Subsidy to wage earners $EITC$. This is so even if $g_0 > g_i$
- Intuition: Increasing the transfer to those without work $(-T_0 \uparrow)$ has a negative incentive effect since it makes not working more attractive for all those who currently work - increasing the transfer to those who work with a low wage has no such negative incentive effect (no intensive margin here).
- Reducing the tax (increasing the transfer) for group 1 with a small amount $(-dT_1 = dc_1 > 0)$. Has three effects
  - Mechanical effect on government income: $-h_1 dT_1$
  - Increased consumption for group 1: $g_1 h_1 dT_1$
  - Behavioral effect gives a tax loss $-(T_1 - T_0)dh_1$ with $dh_1 = -h_1 \cdot \eta_i dT_i/(c_1 - c_0)$ gives a tax loss
- Adding the effects and set them equal to 0 gives the formula for optimal transfers/taxation.

# Extensive and intensive margin: Saez 2002.



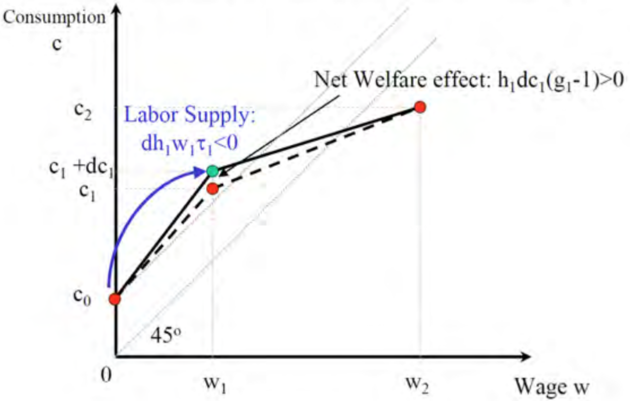
# Extensive and intensive margin: Saez 2002.

Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1 > 1$)

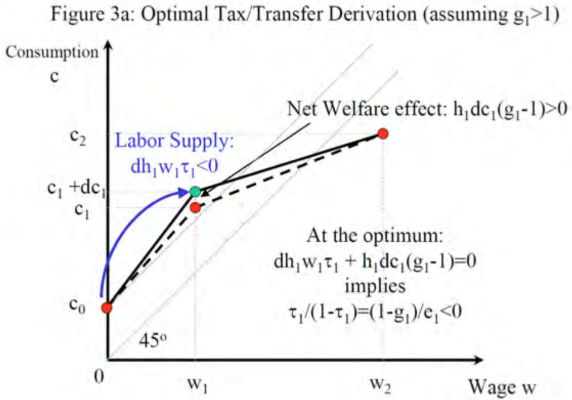


# Extensive and intensive margin: Saez 2002.

Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1>1$)



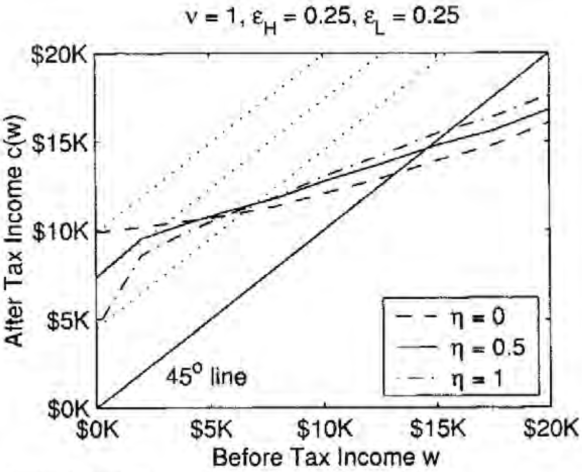
# Extensive and intensive margin: Saez 2002.



# Extensive and intensive margin: Saez 2002.

- Cannot assume away intensive margin, the fact that individuals also decide to work less (not zero) - earn less income - if their income class is taxed more heavily.
- If we include both the intensive and extensive margin it is more complicated: increasing $c_1$ (decreasing $T_1$) has a positive extensive margin effect (individuals will go from 0 wage to ). But it will also reduce others income since some will move from $w_2$ to $w_1$.
- Very useful simulations in the Saez (2002) paper.

# Extensive and intensive margin: Saez 2002.




Source: Saez 2002

# Optimal tax rates at the top income bracket

## Mike Brewer; Emmanuel Saez; Andrew Shephard: Means-testing and Tax Rates on Earnings

First a digression (a useful one): *The laffer curve*.

Assume pretax income is given $y$ and there is a linear income tax (same marginal tax rate for all income). The earned (reported) income will then be a function of the net-of-tax-rate $(1-\tau)$; $y(1-\tau)$.

The tax revenue for the government is $TR(\tau) = \tau \cdot y(1-\tau)$. Suppose the government wants to maximize $TR(\tau)$: Inverse $U$: interior max.

F.o.c. $TR'(\tau^*) = 0 \Rightarrow y - \tau^* \frac{dy}{d(1-\tau)} = 0 \Rightarrow \tau^* = \frac{1}{1+\varepsilon}$ where $\varepsilon = \frac{dy}{d(1-\tau)} \frac{(1-\tau}{y}$. The more elastic taxable income is the lower is the revenue maximizing. Strictly inefficient to have a higher tax rate than $\tau^*$.

# Optimal tax rates at the top income bracket

- Most tax systems have marginal tax rates that are piecewise linear. How should the government decide the tax rate for different segments?
- Changing the tax rate within one income bracket has two effects, it changes the revenue for the government and it changes the welfare, the utility, of the tax payers who have an income within that bracket or above - those with an income below are not affected.
- It is simplest to analyses a change in the top bracket, since a change here will only affect one group; the top earners. If the government were to change the tax rate of a bracket further down in the income distribution it would also affect the average tax rate of those who earn an income above this bracket

# Optimal tax rates at the top income bracket

$N$ individuals earn more than $\bar{y}$ ( lower threshold for the top bracket of the income). Income above this level is at a marginal tax rate of $\tau$.

Average income of those who earn above $y^*$ is equal to $y^m$, this average income will be a function of $(1 - \tau)$, the net of tax income within this bracket.

Consider a small increase in the marginal tax rate $(d\tau)$ of top earners.

For simplicity assume away all income effects so the elasticity of of labour supply is the compensated elasticity, denoted $\bar{\varepsilon}$.

There will be adjustment of income within this bracket, none of the top income earners will reduce their income below $\bar{y}$ since the tax rate there is not changed. Draw.

# Optimal tax rates at the top income bracket

Government Revenue effect ($dR$):

- We can decompose the Revenue effect into a mechanical effect ($dM$) and a behavioral effect ($dB$): $dR = dM + dB$.
- $dM = d\tau\,(y^m - \bar{y})\,N$.
- To find $dB$ note that $y^m$ is a function of the net-of-tax rate. We have $dB = N\tau dy^m = -N\tau \frac{dy^m}{d(1-\tau)} d\tau = -N \frac{\tau}{(1-\tau)} \bar{\varepsilon} y^m d\tau$

# Optimal tax rates at the top income bracket

The welfare effect ($dW$) has two terms.

1. Change in utility of the individuals who are in this income bracket and
2. The weight these individuals are given in the social welfare function.

Since we assume that top earners have chosen their labour supply optimally given the marginal tax rate ($\tau$), we know, from the envelope theorem, that the value of their loss (measured in NOK) is equal to $dM$. Suppose that the government assign a weight $\bar{g} \in (0, 1)$ on this loss (if it is equal to 0 the income lost at the top does not have any effect on social welfare; in general $\bar{g}$ depends on the marginal utility of income for top earners and on the shape of the social welfare function: $dW = -\bar{g}\,(d\tau\,(y^m - \bar{y})\,N)$.

# Optimal tax rates at the top income bracket

- Adding together these terms gives the total social welfare effect of the small increase in the tax rate ($d\tau$).
  - The initial tax rate will be optimal if the total effect of increasing the rate is 0. Hence to characterize the optimum top income marginal tax rate ($\bar{\tau}$) we solve $dM + dB + dW = 0$. We get the following expression

$$\frac{\bar{\tau}*}{1-\bar{\tau}*} = \frac{(1-\bar{g})(\frac{y^m}{\bar{y}}-1)}{\bar{\varepsilon}\frac{y^m}{\bar{y}}}$$

  - The optimal tax rate of top earners
    - decreases in $\bar{g}$; the more weight the loss of the top earners is given
    - decreases in the compensated elasticity of labour supply
    - increases in the thickness of the tale of the income distribution in the top bracket ($\frac{y^m}{\bar{y}}$)
    - if $\frac{y^m}{\bar{y}} = 2, \bar{g} = 0,5, \bar{\varepsilon} = 0,1 \Rightarrow \bar{\tau}* = 71\%$