Seminar on Public Finance

Lecture #4: February 6

Optimal Taxation

Optimal Taxation

- What do we mean by optimal?
- According to Salanié there are 6 criteria (the first four noted by Adam Smith in the *Wealth of Nations*:
  1. Taxes must be related to the ability to pay
     - Horizontal versus vertical equity
  2. Taxes must be clearly defined and not arbitrary
  3. Taxes must be collected in a painless manner
     - Low compliance costs
  4. Taxes must have low costs
     - Low administrative costs and low efficiency costs
  5. Taxes must adapt to economic fluctuations as automatic stabilizers
  6. Tax incidence should be clear

Optimal Taxation: What's the problem?

- Goal is to maximize social welfare (minimize DWL) subject to a revenue constraint
- First best:
  - Suppose we have perfect information, complete markets, perfect competition, lump sum taxes feasible at no cost
  - Result: Second welfare theorem implies that any Pareto-efficient allocation can be achieved as a competitive equilibrium with appropriate lump-sum transfers (or taxes)
  - Economic policy problem reduces to the computation of the lump-sum taxes necessary to reach the desired equilibrium.
  - The equity-efficiency tradeoff disappears.

Optimal Taxation: What's the problem?

- Problems with the first best:
    - We don't really see lump sum taxes in the real world!
    - No way to make people reveal their characteristics at no cost
        - e.g., if going to redistribute from those with high ability to those with low ability, a skilled person would pretend to have low ability
    - Thus, the optimal taxation literature rules out the possibility of lump-sum taxation
    - So gov't has to set taxes as a function of economic outcomes: receipt income, consumption $\Rightarrow$ distortion and DWL
    - A tradeoff between equity and efficiency in the 2nd best.

Optimal Taxation (2)

- By taxing economic transactions we potentially alter the economic decisions of individuals and thus create inefficiencies
- Thus the problem that the optimal taxation literature studies is how much tax should be levied on each good or factor income to raise a given amount of government revenue to maximize social welfare
    - As we have discussed previously, defining a social welfare function is non-trivial
    - Fortunately, most of the basic results dont rely on a specific function form

Optimal Taxation (3)

- There are four main results in optimal tax theory:
    1. Ramsey inverse elasticity rule
    2. Mirrlees taxation: production efficiency
    3. Atkinson and Stiglitz: no consumption taxation with optimal non-linear (including lump-sum) income taxation [not covered here]
    4. Chamley/Judd: no capital taxation in infinite horizon models [not covered here]

Optimal Taxation (3)

- There are 3 main areas of optimal taxation:
  1. Optimal commodity taxation
     - Optimal linear taxes on goods and services
  2. Optimal income taxation
     - Optimal non-linear taxes on income
  3. Optimal mixed taxation
     - Combines linear consumption taxes with non-linear income taxes

Optimal Commodity Taxation

- Commodity taxation is often termed indirect taxation since the tax is independent from the person who bears the burden.

- The problem in the theory of optimal commodity taxation is: What is the appropriate policy for setting commodity tax rates?

- The objective is to raise some given amount of revenue in a way that minimizes the economic cost of collecting the revenue.

  - Taxes are chosen to maximize the utility of the residents, or equivalently, minimize excess burden, subject to revenue constraint

- Note that by design this is considering the case where income is exogenous.

- The taxes are restricted to be linear taxes

Optimal Commodity Taxation (2)

- Assume two commodities, X and Y, and leisure, L.
- Price of X is P_x, price of Y is P_y, and the price of leisure is the wage rate, w.
- The maximum hours that can be worked per week is the time endowment, $\bar{T}$, which is fixed. Typically, this is considered 24 hrs per day.
- The difference $\bar{T} - L$ is the hours worked per week. Substituting these into the budget constraint, $I = P_x X + P_y Y$, yields:

$$w(\bar{T} - L) = P_x X + P_y Y$$

- The LHS gives total earnings; RHS shows total consumption.

Optimal Commodity Taxation (3)

- Rearranging, we obtain:

$$w\bar{T} = P_x X + P_y Y + wL$$

- This equation shows the total (maximum) earnings possible if an individual worked 24/7.
- The LHS, $w\bar{T}$, is referred to as the value of the time endowment.
- Now assume we can impose ad valorem taxes on each of the two commodities *and* leisure at equal rates

Optimal Commodity Taxation (4)

- The budget constraint becomes:

$$w\bar{T} = (1+t)P_x X + (1+t)P_y Y + (1+t)wL$$

- Simplifying this expression yields:

$$\frac{w\bar{T}}{(1+t)} = P_x X + P_y Y + wL$$

- Hence an ad valorem tax of equal value placed on all commodities, *including leisure*, generates a lump sum tax, which is non-distortionary and produces no excess burden.

- However, this is not feasible since it is not possible to tax leisure.

- Also is the wage rate truly exogenous?

Optimal Commodity Taxation (5)

Whats the solution? An obvious first guess:

- Apply the same rates of taxation across all goods i.e. tax all goods at 5%, for example, otherwise known as *neutral or uniform taxation*.
- Consider this case shown below. Goods are X and Y, but each have different elasticities. The price for each is the same before and after tax (for simplicity).
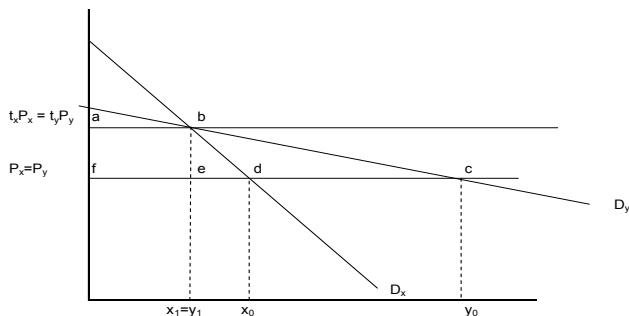
Optimal Commodity Taxation (6)

- The amount of tax revenue collected from the tax is the same for each good: area $abef$.
- However, due to the different demand elasticities, the excess burdens, EB, are very different:
  - $EB_X = bde$
  - $EB_Y = bce$
- Thus, neutral taxation tends to be very distortionary.

- A second guess: Just raise existing rates on all goods to increase tax revenues - double them, for example.

Optimal Commodity Taxation (7)
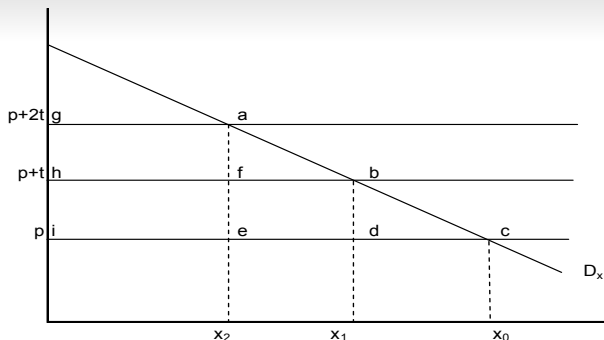
- TR with tax t is the area $hbdi$; EB with tax t is bcd.
- TR with tax $2t$ is the area $gaei$; EB with tax $2t$ is ace.
- Excess burden with tax $2t$ is 4 times that with tax t.
  - The area of the top triangle equals the area of the bottom triangle and the area of $fbde$ is two times the area of the top triangle hence ace is 4 times the area of bcd, even though the tax rate has only doubled.

Optimal Commodity Taxation (8)

- This is known as the square rule.
    - When both supply and demand curves are linear, increasing a tax rate will increase the excess burden at an *increasing rate*.
- We can now make two conclusions. The amount of excess burden from a single tax depends on:
    - the magnitude of the tax
    - the elasticity of demand

Optimal Commodity Taxation (9)

Whats the solution? *The Ramsey Rule*

- Given that different demand elasticities are present, to minimize overall excess burden across all *unrelated* commodities, the marginal excess burden of the last dollar of revenue raised from each commodity must be equal.

$$\frac{MEB_X}{MR_X} = \frac{MEB_Y}{MR_Y}$$

- This is an "equal marginal rule" (or arbitrage rule) like we typically find in micro theory.

- Back to our examples of unrelated good X & Y
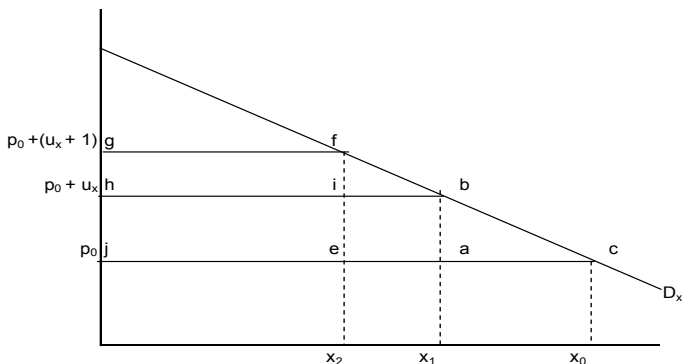
- Our initial tax is u_X and then we add a small (marginal) increment to it.

Optimal Commodity Taxation (11)

- With our initial tax in place the excess burden is $0.5u_X(x_1 - x_0)$ and total revenue is $u_X x_1$
- Now we calculate the marginal excess burden:
  - New $EB = 0.5(u_X + 1)(x_2 - x_0)$
  - $MEB = 0.5(u_X + 1)(x_2 - x_0) - 0.5u_X(x_1 - x_0) \approx 0.5(x_2 - x_0)$ as tax change goes to zero
- The marginal revenue from the tax change:
  - $MR = (u_X + 1)x_2 - u_X x_1 \approx x_2$ as tax change goes to zero
- In a similar fashion find the MEB and MR from a incremental tax on commodity y

Optimal Commodity Taxation (12)

- Putting this together we see that:

$$\frac{MEB_X}{MR_X} = \frac{0.5(x_2 - x_0)}{x_2} = \frac{\Delta x}{x_{new}}$$

$$\frac{MEB_Y}{MR_Y} = \frac{0.5(y_2 - y_0)}{y_2} = \frac{\Delta y}{y_{new}}$$

$$\frac{MEB_X}{MR_X} = \frac{\Delta x}{x_{new}} = \frac{\Delta y}{y_{new}} = \frac{MEB_Y}{MR_Y}$$

- This result is the Ramsey Rule: To minimize total excess burden, tax rates should be set so that the percentage reduction in the quantity demanded of each good is the same.

- Note: this implies that the tax rates need not be equal

Optimal Commodity Taxation (13)

- The Ramsey Rule is: $\frac{MDWL_i}{MR_i} = \lambda \Rightarrow \tau = \frac{\lambda}{\eta_D}$
    - Note that the Ramsey Rule here is written in the case of infinitely-elastic supply - so all incidence on consumers, hence only the demand elasticity, η_D above
- It sets taxes across commodities so that the ratio of the marginal deadweight loss to marginal revenue raised is equal across commodities.
    - The goal of the Ramsey Rule is to minimize deadweight loss of a tax system while raising a fixed amount of revenue.
    - λ measures the value of having another dollar in the government's hands relative to the next best use in the private sector.
        - Smaller values of λ mean additional government revenues have little value relative to the value in the private market.

Derivation of the Inverse Elasticity Rule

- More formally, consider the excess burden from imposing a small ad valorem tax t_i on good i:

$$DWL_i(t_i) = \frac{\eta_S^i \eta_D^i}{\eta_S^i + \eta_D^i} t_i^2 \frac{p_i x_i}{2}$$

  - Note that I used a sum in the denominator- the reason is to clarify that we are adding magnitudes of elasticities (recall that the demand elasticity is negative, so here we are just adding the negative of the demand elasticity instead of subtracting the (negative) demand elasticity)

- The revenue generated by a vector of n tax rates on goods 1 to n, is:

$$R(t) = t_1 p_1 x_1 + ... + t_n p_n x_n$$

- The total excess burden in partial equilibrium is:

$$DWL(t) = DWL_1(t_1) + ... + DWL_n(t_n)$$

Derivation of the Inverse Elasticity Rule(2)

- The optimal problem is to raise a given about of revenue $R(t) = T$ while minimizing $DWL(t)$

$$\mathscr{L} = \sum_{i=1}^{n} \frac{\eta_S^i \eta_D^i}{\eta_S^i + \eta_D^i} t_i^2 \frac{p_i x_i}{2} + \lambda \left(T - \sum_{i=1}^{n} t_i p_i x_i \right)$$

- In partial equilibrium the expenditure on each good is fixed, so $p_i x_i$ is a constant.

- Each of the first n first order conditions imply that:

$$\frac{\eta_S^i \eta_D^i}{\eta_S^i + \eta_D^i} t_i = \lambda \Rightarrow t_i = \lambda \left(\frac{1}{\eta_S^i} + \frac{1}{\eta_D^i} \right)$$

- This is the "inverse elasticity rule"

Derivation of the Inverse Elasticity Rule(3)

- The inverse elasticity rule, based on the Ramsey result, allows us to relate tax policy to the elasticities.
    - The government should set taxes on each commodity inversely to the total elasticity.
    - Therefore, ignoring equity, less elastic items should be taxed at a higher rate.
- Two factors must be balanced when setting optimal (efficient) commodity tax rates (again, ignoring equity):
    - *The elasticity rule*: Tax commodities with low elasticities.
    - *The broad base rule*: It is better to tax many goods at lower rates, because DWL increases with the square of the tax rate.
        - Thus the government should tax all of the commodities that it is able to, but at different rates.

"Problems" with the Ramsey Rule

- Following this rule would imply that:
  - Impose high tax rates on insulin, AIDS drugs, food, ...
  - Impose low taxes on luxury goods, yachts, ...
- The prescription clearly violates equity principles which in hindsight is unsurprising since equity wasn't explicitly considered.
- Often there is a high correlation between the degree of inelasticity and the utility cost.
  - Necessities are inelastic
  - Luxuries are elastic

Implication of the Ramsey Rule: Corlett-Hague Rule

- Recall that in a first-best situation, leisure would be taxed. However, it is not possible to directly tax leisure.
- As a result an interesting implication of Ramsey Rule is that:
  - In a two commodity world, efficient taxation requires that the commodity that is complementary to leisure be taxed at a relatively high rate: the Corlett-Hauge Rule
- By taxing good jointly consumed with leisure, the demand for leisure can be lowered, and we can move closer to a more efficient outcome.

Optimal Income Taxes

- Optimal income taxation is choosing the tax rates across income groups to maximize social welfare subject to a government revenue requirement.
- Simple models lead to the conclusion that top rates would be 100% - e.g. models with exogenous labor supply
- Raising tax rates will likely affect the size of the tax base. Thus, increasing the tax rate on labor income has two effects:
  - Tax revenues rise for a given level of labor income.
  - Workers reduce their earnings, shrinking the tax base.
- At high tax rates, this second effect becomes important. Which leads to the famous Laffer curve
  - Thus, there are equity-efficiency tradeoffs in designing income tax rates.

The *Laffer curve*
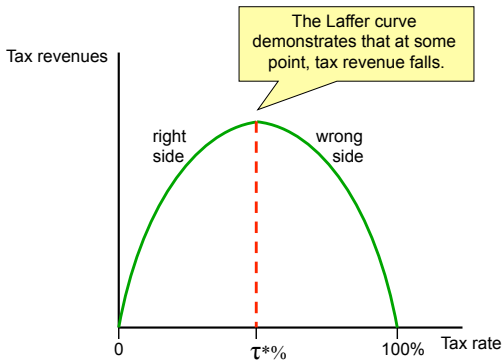
Optimal income taxes: A general intuitive model with behavioral effects(2)

- The goal of optimal income tax analysis is to identify a tax schedule that maximizes social welfare, while recognizing that raising taxes has conflicting effects on revenue.
- The optimal tax system meets the condition that tax rates are set across groups such that: $\frac{MU_i}{MR_i} = \lambda$
    - Where MU_i is the marginal utility of consumption for individual i, and MR_i is the marginal revenue from that individual.
- As with optimal commodity taxation, this outcome represents a compromise between two considerations:
    1. Vertical equity
    2. Behavior responses

Optimal income taxes: A general intuitive model with behavioral effects(3)
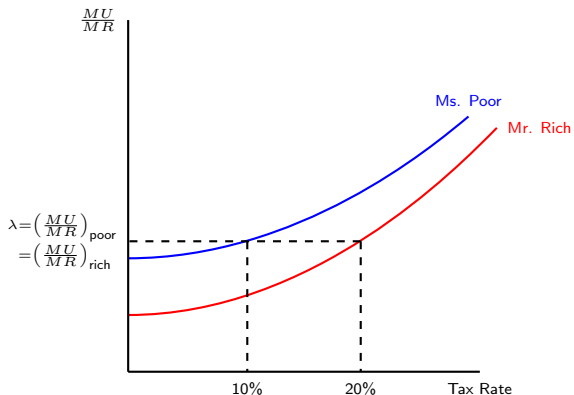
Low income may have higher MU_c. A given tax will raise more for a high income household (up to a point), so MR is higher.

Mirrlees Model (Simplified)

- Assume an unequal distribution of income and that the tax affects the labor supply decision.

- Consumers have identical preferences but differing skill levels and hence wage.

- It is a competitive economy so workers are paid their marginal product and firms price output at marginal cost.

- The government can only tax income (the product of wage and hours worked.)

- Income tax schedule is chosen to maximize social welfare subject to raising a given amount of revenue.

Mirrlees Model (2)

- There are two goods: consumption, X, and leisure, H.
- Let $z = wL$ where L is labor hours.
- Let $T(z)$ be the amount of tax paid on earnings z, hence $x = c(z) = z - T(z)$
- The optimization problem will determine the shape of the function $T(z)$.
  - This is a difficult optimization problem where the result depends heavily the parameters assumed.
  - The following graph depicts a generalization of what this function may look like given a typical set of parameters.

Mirrlees Model (3)

- Without a tax $x = z$ so the tax generates a departure from the $45°$ line.

- Where the consumption function, c, lies above the $45°$ line there is a payment to the individual (i.e., a negative income tax).

- The slope of the consumption function is 1-(marginal tax rate)
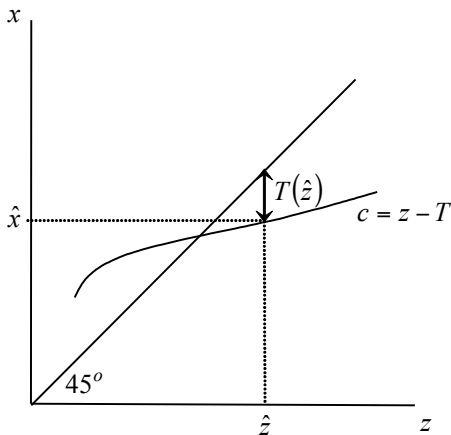
Mirrlees Model (4)

- Since the utility functions are identical but depend on skill levels, they will only cross once.

- Thus high skill workers will always earn at least as much a low skill workers.

- Workers choose how much to work by finding the tangency between the indifference curve and the consumption function.
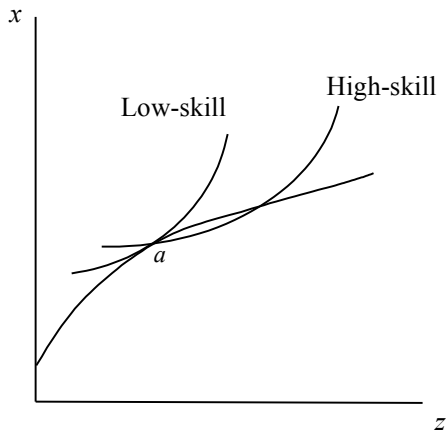
Mirrlees Model (5)

- Since $c(z) = z - T(z), c'(z) = 1 - T'(z)$
- But $c'(z) \geq 0$ (must get something for additional effort) which implies that $T'(z) \leq 1$, the marginal tax rate must be less than 100%.
- Is there a minimum tax rate? Might the marginal rate be negative $(T'(z) < 0)$?
    - A negative tax rate would imply that the slope of the consumption function was greater than 1 (the slope of the $45°$ line).
    - It can be shown that a the marginal tax rate will never be negative.
    - Think about negative marginal rate as incentive to earning more income (the more you earn, the larger the tax refund)
    - But high skill must work less hours for given amount of income, so increases his utility the most relative to zero MTR
        - Thus, can increase rate and have revenue gain, while lowering utility of high skill less than increase utility of low skill

Mirrlees Model (6)

- As we move from $c_1(z)$ to $c_2(z)$ the change between the high skill and low skill worker offset.

- This change must raise welfare since the MU is higher for the low skill worker than the high skill.

- Thus the marginal tax rate must be non-negative, $T(z) \geq 0$

- Note that average tax rates may be negative (e.g. phase out benefits for poor)
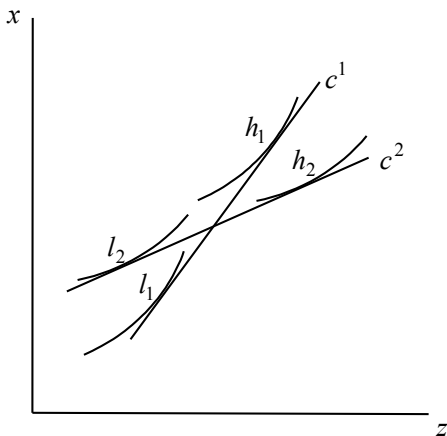
Mirrlees Model (7)

- What is the optimal marginal rate for the highest skill worker?

- Consider the consumption function ABC where the worker locates at B.

- A clearly superior function is ABD.
    - Segment BD is at $45°$.
    - Worker is better off at point b and revenue is unchanged.

- The optimal tax function must have a top marginal rate of 0%.
    - This result turns out not to be a general one.
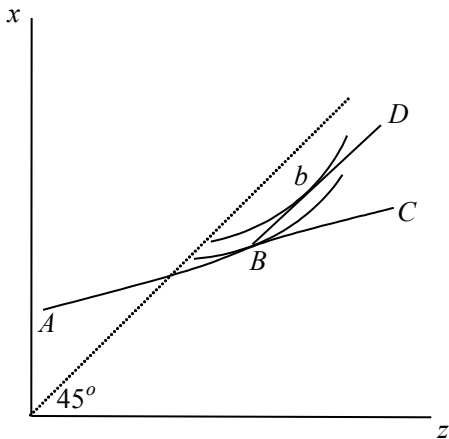
Mirrlees Model (8)

- The basic results of Mirrlees is that:
  - Marginal tax rates are non-negative
  - Marginal tax rates should not exceed 1
  - The marginal tax rate facing the highest skilled (paid) worker should be zero
  - While formally non-linear the optimal tax is approximately linear
- Others have found that if everyone works the lowest skilled worker should also face a zero rate.
- The actual shape of the tax function depends critically on the shape of the skill/wage distribution.
- The zero top rate result only applies to the very highest taxpayer and hence is not considered a meaningful result.

Saez (2001)

- Saez (2001) repeats Mirrlees, but describes optimal tax structure in terms of elasticities and parameters of the income distribution.
- He finds that you want a low tax rate when:
    - The elasticity of labor supply large
    - The density of the wage distribution is high
    - At the top of the distribution
- In addition, the SWF used has large implications on the level of rates (but not so much on the general shape of the marginal rate structure)
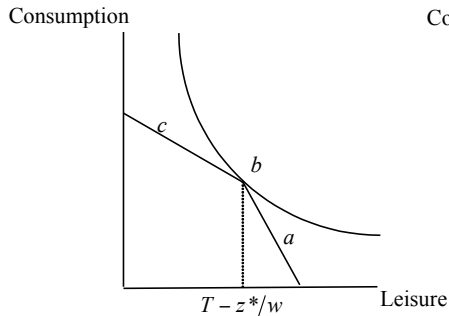
Simulations by Saez (2001)
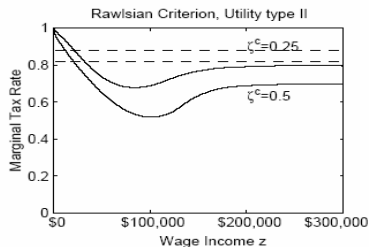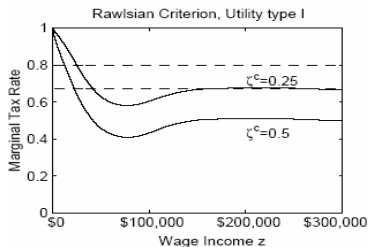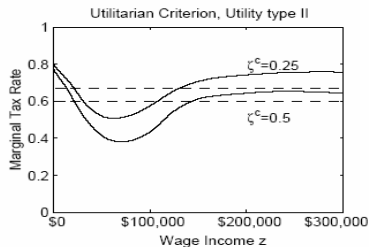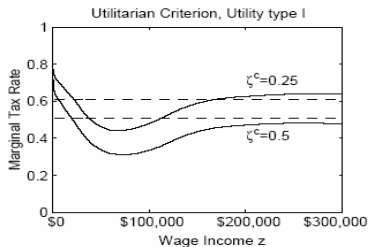
Simulation by Mankiw et al
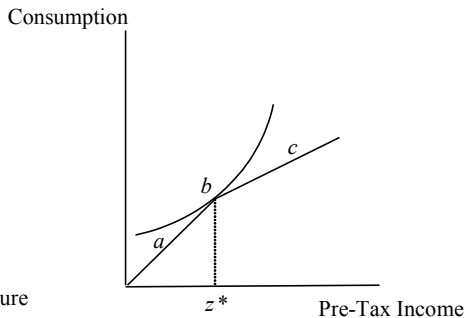

Figure 3: Optimal Marginal Tax Simulations
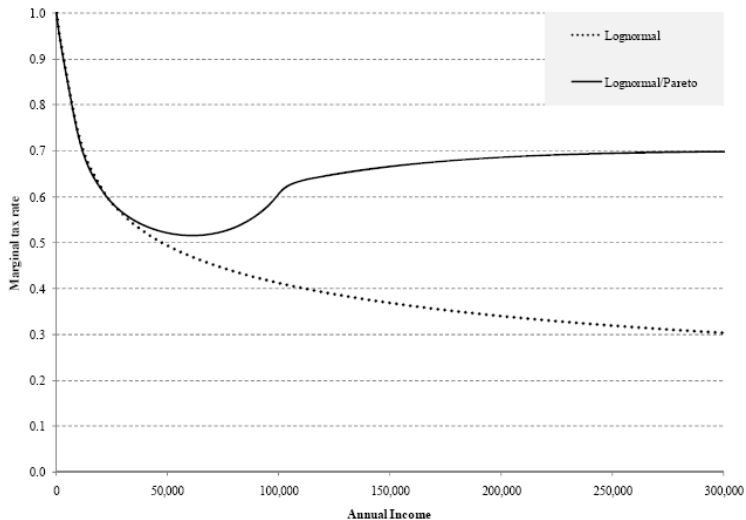
Progressive Taxation of Income

- In the real world labor is taxed under a progressive rate schedule.
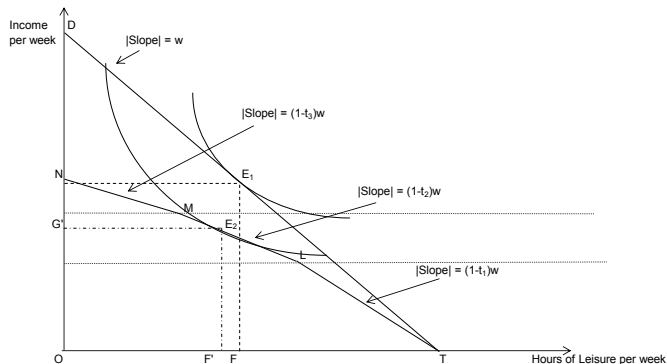- This introduces kink points where rates change.
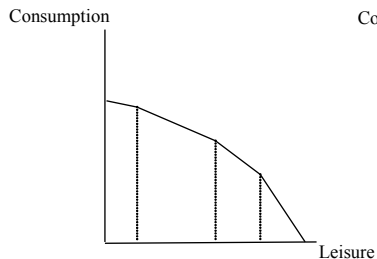

a. Leisure

b. Pre-Tax Income

Progressive Taxation of Income (2)

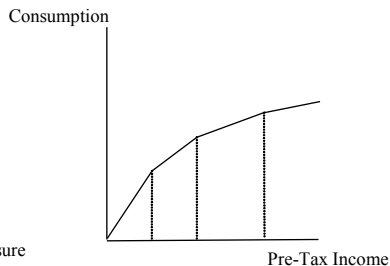- We have one less kink than we have rate brackets

Progressive Taxation of Income (3)

- With multiple rate brackets we get a budget constraint that is "bowed out."
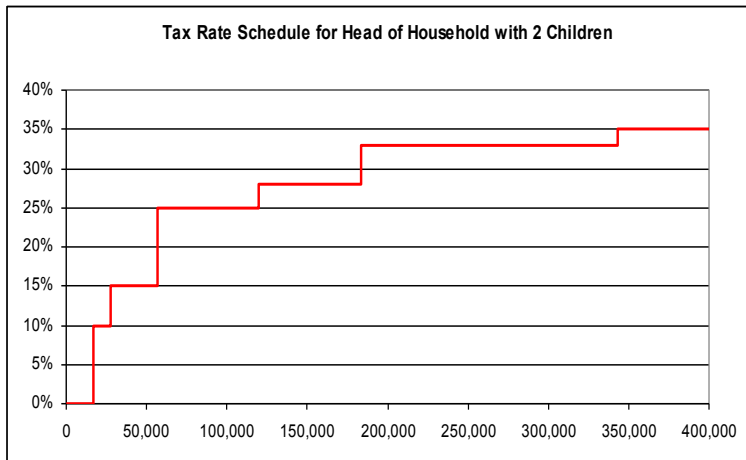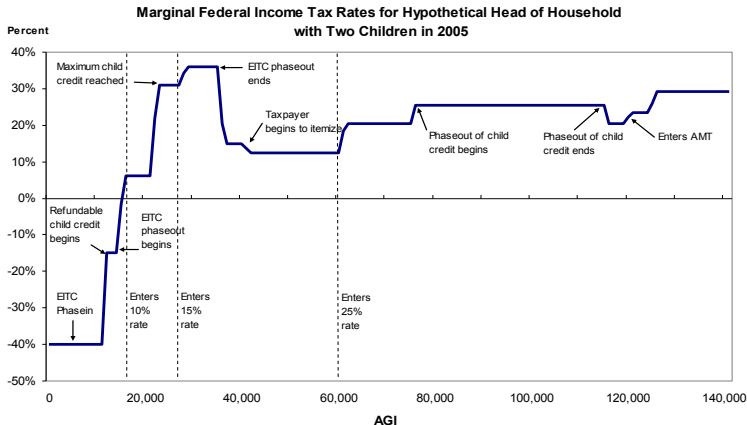

a. Leisure

b. Pre-Tax Income

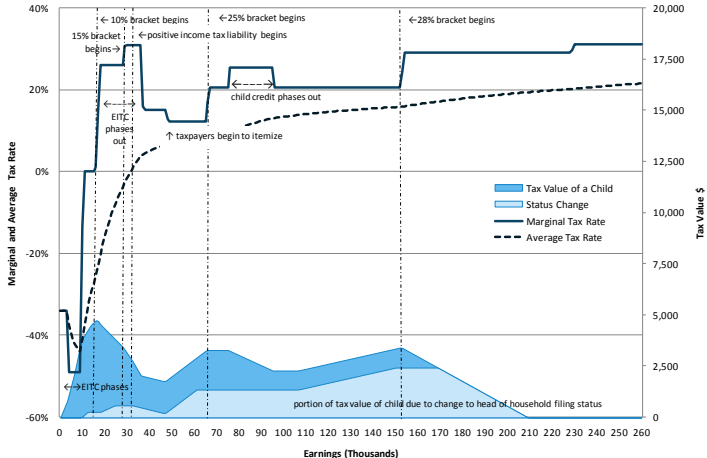# Progressive Taxation of Income (4): Marginal Tax Rates in "Theory"

Tax Rate Schedule for Head of Household with 2 Children

Progressive Taxation of Income (5): Marginal Tax Rates in "Reality"

**Marginal Federal Income Tax Rates for Hypothetical Head of Household with Two Children in 2005**

Note: Calculations are for a head of household with two children under 17. Itemized deductions are assumed to be 18 percent of income.

Figure 5
Tax Rates and Tax Value of a Child for HoH Filer with One Dependent Child
(2011 Law)

Progressive Taxation of Income (7): From Kotlikoff and Rapson 2006, Impact of "all" government programs

Marginal Effective Tax Rate
30 Year Old Couples Earning $0-$50,000/yr
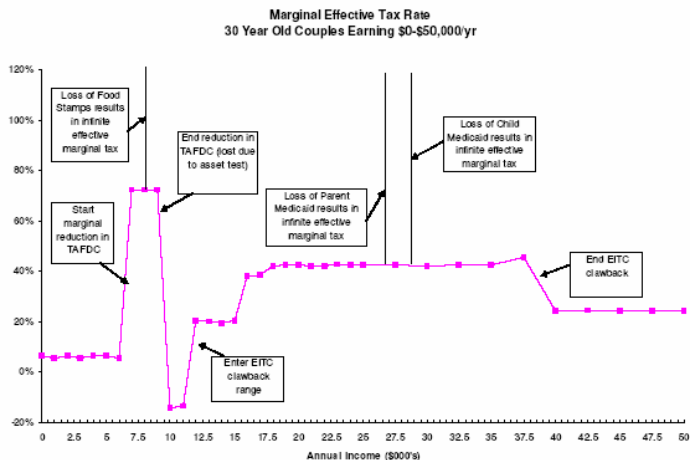
Optimal Tax in Theory and Practice: Mankiw, Weinzierl and Yagan

1. Optimal marginal tax rate schedules depend on the distribution of ability

- This takes on added importance as the returns to education have seemed to have increased
- Should we take wages as a proxy for ability?
- Income distribution has become much more skewed.

Optimal Tax in Theory and Practice (2): Mankiw, Weinzierl and Yagan

2. The optimal marginal tax schedule could decline at high incomes

- The degree that this is true depends on the taxable income elasticities
- The empirical literature finds that the taxable income elasticities increase with income
- Don't see this in the labor elasticity
- Is it a real effect or simply a timing or compliance effect?

Optimal Tax in Theory and Practice (3): Mankiw, Weinzierl and Yagan

3. A flat tax, with a universal lump-sum transfer, could be close to optimal

- This results in large differences in average and marginal tax rates
- Do people respond differently? Look to the new literature on tax salience
- How should we think of non-tax increases in the social safety net, are these lump sum?

Optimal Tax in Theory and Practice (4): Mankiw, Weinzierl and Yagan

4. The optimal extent of redistribution rises with wage inequality
   - Clearly inequality is increasing
   - There is some evidence that incremental changes to tax systems around the world are responding this change

Impact of the change in the wage distribution

Figure 4: Optimal average tax rates, 1979 and 2007
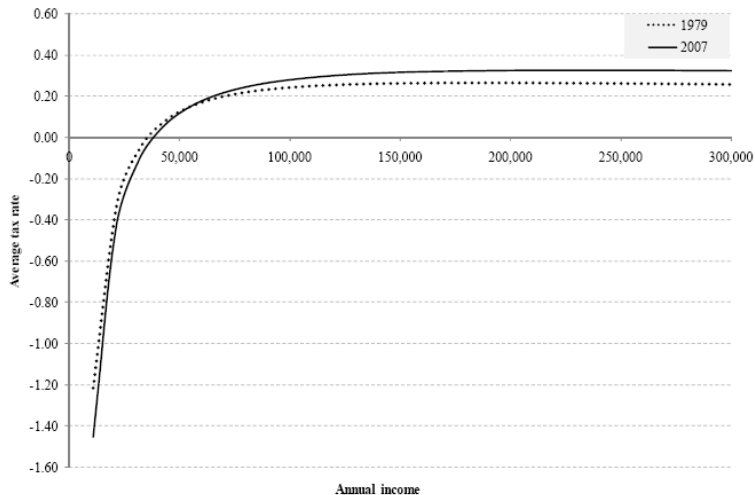
Optimal Tax in Theory and Practice (5): Mankiw, Weinzierl and Yagan

5. Optimal taxes should depend on personal characteristics as well as income

- Theoretically tax should depend on ability - problem is that we can't observe this directly
- Current tax system does rely on some family characteristics (e.g. married, children)
- This is an area where economists make a prescription that is not widely acceptable to those outside the profession

Optimal Tax in Theory and Practice (6): Mankiw, Weinzierl and Yagan

6. Only final goods ought to be taxed, and typically they ought to be taxed uniformly

- This is a strong result in the literature on indirect taxes and seems most directly applicable to value added taxes
- In practice both VATs and sales taxes have lots of exemptions which greatly reduce both the revenue raised and the efficiency of these indirect taxes
- Exception to this rule on uniform taxes is externalities

Optimal Tax in Theory and Practice (7): Mankiw, Weinzierl and Yagan

7. Capital income ought to be untaxed, at least in expectation

- General intuition is that this is a tax on saving and as a result it distorts the timing of consumption. Thus it would be superior to directly tax consumption rather than income and savings

- Note that the models presume that savings and capital is very responsive to taxation, if it was less responsive then it would be less distortionary

- Why do people save? To consume later? Precautionary reasons?

- Progressivity of most capital taxes

Optimal Tax in Theory and Practice (8): Mankiw, Weinzierl and Yagan

8. In stochastic, dynamic economies, optimal tax policy requires increased sophistication

- Most models dont handle lifecycle effects well
- Interaction between labor taxes and capital taxes need to be considered
- Over what time frame should we measure the optimal tax given intertemporal and even intergenerational effects?