14.471 Notes on Linear Taxation

Iván Werning

1 Overview

- Two models
  - single agent (Ramsey), no lump sum tax
  - agent heterogeneity and lump sum tax
- Two approaches
  - primal
  - dual
- Mixed Taxation

2 Single Agent Ramsey

- consumers:
  \[
  \max_x u(x) \quad \sum_i q_i x_i \leq 0
  \]
  e.g. \( u(c_1, c_2, \ldots, c_n, l) \) and \( \sum p_i (1 + \tau_i) c_i = (1 - \tau^l) w l \)
- CRS technology (inputs are supressed)
  \[
  F(y) \leq 0
  \]
  e.g. \( \sum \bar{p}_i y_i - l \leq 0 \)
- Remark: Production efficiency holds so that \( F(y) = 0 \) at optimum
  (implies intermediate inputs go untaxed)
  without CRS this result requires profit taxes (see Diamond-Mirrlees)
• First Best
  - \( MRS_{ij}^h = MRS_{ij}^h \)
  - \( MRS_{ij}^h = MRT_{ij} \)
  - \( F = 0 \) (efficient production; with inputs this requires a marginal condition equating the relative marginal products across goods)

• Firms
  \[
  \max_Y p_y \quad F(y) \leq 0
  \]

• government
  \[
  \sum p_i g_i \leq \sum t_i x_i
  \]

• market clearing:
  \[
  x_i + g_i = y_i \quad \forall i
  \]

• note: we could have \( u(c, g) \), but in what follows \( g \) is fixed, so we suppress the dependence.

• Definition: A Competitive Equilibrium (CE) with taxes is \( p, q, c \)
  
  1. \( x \) solves the consumer’s maximization
     \[
     \max_x u(x) \quad \sum_i q_i x_i \leq 0
     \]

  2. \( y \) solves the profit maximization
     \[
     \max_y p_y \quad F(y) \leq 0
     \]

  3. \( x, g, t, p \) satisfy the government budget constraint
     \[
     \sum p_i g_i \leq \sum t_i x_i
     \]

  4. markets clear
     \[
     x_i + g_i = y_i \quad \forall i
     \]

• Result: CE \( \iff \) \( F(x + g) = 0 \) and agent optimization (1)

• note: second condition involves \( x \) and \( q \) only
• First Best

\[
\max_{x,q} u(x) \\
F(x + g) = 0
\]

• Second Best

\[
\max_{x,q} u(x) \\
F(x + g) = 0
\]

\[x \in \arg\max_x u(x) \quad q \cdot x \leq 0\]

• we have two variables \(x, q\) but they are related through the last condition

• At this point, from consumer maximization we can approach things from...
  
  – primal: solve \(q\) as a function of \(x\)
  
  – dual: solve \(x\) as a function of \(q\)

• both approaches are useful

2.1 Dual

• define

\[
V(q, I) = \max_x u(x) \quad q \cdot x \leq I
\]

and let \(x_i(q, I)\) denote the solution (Marshallian/uncompensated demand)

\[
e(q, v) \equiv \min_x q \cdot x \quad u(x) = v
\]

and let \(x^c_i(q, v) = e_{qi}(q, v)\) denote the solution (Hicks/compensated demand)

• we abuse notation: \(V(q) = V(q, 0)\)

• Second Best:

\[
\max_q V(q) \quad \text{s.t.} \quad F(x(q, 0) + g) = 0
\]

• property:

\[
x^c(q, V(q)) = x(q, 0)
\]
• equivalently
\[ \max_q V(q) \quad \text{s.t.} \quad F(x^c(q, V(q) + g) = 0 \]

2.2 Optimality condition

• We have the first order condition
\[ \frac{\partial V}{\partial q_j}(q, 0) - \kappa \sum_i \frac{\partial F}{\partial y_i} \left( \frac{\partial x_i^c}{\partial q_j} + \frac{\partial x_i^c}{\partial v} \frac{\partial V}{\partial q_j} \right) = 0 \]

• By Roy's identity \( \frac{\partial V}{\partial q_j} = -x_j \frac{\partial V}{\partial I} \):
\[ -\frac{1}{\kappa} x_j \frac{\partial V}{\partial I} - \sum_i \frac{\partial F}{\partial y_i} \left( \frac{\partial x_i^c}{\partial q_j} - x_j \frac{\partial x_i^c}{\partial v} \frac{\partial V}{\partial I} \right) = 0 \]

• Now use that \( \frac{\partial x_i^c}{\partial v} \frac{\partial V}{\partial q_j} = \frac{\partial x_i^c}{\partial q_j} \) and \( p_i = \frac{\partial F}{\partial y_i} \) to get
\[ -\frac{1}{\kappa} x_j \frac{\partial V}{\partial I} - \sum_i p_i \frac{\partial x_i^c}{\partial q_j} + x_j \sum_i p_i \frac{\partial x_i^c}{\partial q_j} = 0 \]

• Now we know that \( \sum_i q_i \frac{\partial x_i^c}{\partial q_i} = 0 \) and that \( \frac{\partial x_i^c}{\partial q_i} = \frac{\partial x_i^c}{\partial q_j} \) by symmetry so that
\[ -\sum_i p_i \frac{\partial x_i^c}{\partial q_j} = \sum_i t_i \frac{\partial x_i^c}{\partial q_j} \]

• Also, we know that \( \sum_i q_i \frac{\partial x_i^c}{\partial q_i} = 1 \) so that
\[ \sum_i p_i \frac{\partial x_i^c}{\partial I} = 1 - \sum_i t_i \frac{\partial x_i^c}{\partial I} \]

• Thus, we obtain
\[ \sum_i t_i \frac{\partial x_i^c}{\partial q_j} = -x_j \theta \]
\[ \theta \equiv -\frac{1}{\kappa} \frac{\partial V}{\partial I} + 1 - \sum_i t_i \frac{\partial x_i}{\partial I} \]

- or equivalently (using symmetry)

\[ \sum_i t_i \frac{\partial x_i^c}{\partial q_i} = -x_i \theta. \]

- interpretation:

  - each good is “discouraged” by a common percentage \( \theta \), i.e. interpret (falsely) as an estimate of how much good \( x_i \) fell due to taxation.

  - \( \text{DWL} = e(q, V(q)) - \sum t_i x_i^c(p, V(q)) \)

\[ \frac{1}{x_i p_i} \frac{\partial \text{DWL}}{\partial \tau_i} = \text{constant} \]

  intuitive: marginal DWL is proportional to revenue base (mg cost = mg benefit)

2.3 Primal

- Primal solves \( q \) from \( x \)

- consumer optimization

\[ x \in \arg \max_x u(x) \quad q \cdot x \leq 0 \]

- necessary and sufficient conditions: \( \exists \lambda > 0 \) s.t. (assuming local non-satiation)

\[ q_i = \lambda u_i(x) \]

\[ q \cdot x = 0 \]

thus (implementability condition)

\[ \sum u_i(x)x = 0 \]
• Result: reverse is also true: if $\sum u_i(x)x = 0$ then $\exists q$ such that $x \in \arg \max_x u(x)$ s.t. $q \cdot x \leq 0$.

• Second best

$$\max u(x)$$

$$F(x + g) = 0$$

$$\sum u_i(x)x = 0$$

• Lagrangian:

$$L = u(x) + \mu \sum u_i(x)x - \gamma F(x + g)$$

• FOC

$$(1 + \mu)u_i(x) + \mu \sum_j u_{ij}(x)x_j = \gamma F_i(x + g)$$

• implication

$$\frac{F_i(x + g)}{F_k(x + g)} = \frac{u_i(x)}{u_k(x)} \frac{1 + \mu + \mu \sum_j \frac{u_{ij}(x)}{u_i(x)} x_j}{1 + \mu + \mu \sum_j \frac{u_{kj}(x)}{u_k(x)} x_j}$$

• since

$$\frac{F_i(x + g)}{F_k(x + g)} = \frac{p_i}{p_k} \frac{u_i(x)}{u_k(x)} = \frac{q_i}{q_k}$$

• tax rate (where $q_i = \tau_i p_i$)

$$\frac{\tau_i}{\tau_k} = \frac{1 + \mu + \mu \sum_j \frac{u_{ij}(x)}{u_i(x)} x_j}{1 + \mu + \mu \sum_j \frac{u_{kj}(x)}{u_k(x)} x_j}$$

• Exercise: show that if $U(G(x_1, x_2, \ldots, x_n), x_0)$ and $G$ is homogeneous of degree 1 then $\tau_1 = \tau_2 = \cdots = \tau_n$. 
2.4 Many Agents Dual

- Second Best (dual)

$$\max_{q,I} \sum \lambda^h V^h(q, I) \pi^h \quad \text{s.t.} \quad F(\sum x_{c,j}^h(q, V^h(q, I)) \pi^h + g) = 0$$

- note about $I$:
  - we can impose $I = 0$;
  - typically we do not want to: captures a lump sum transfer/tax
  - if we allow $I$ free then productive efficiency is obvious

- more generally
  - Pareto problem not convex
  - cannot maximize weighted utility
  - but pareto weights for local optimality condition

- Define Lagrangian

$$L = \sum \lambda^h V^h(q, I) \pi^h - \gamma F(\sum x_{c,j}^h(q, V^h(q, I)) \pi^h + g)$$

- FOCs: (using same identities as before)

$$- \sum \lambda^h x_j^h \frac{\partial V^h}{\partial I} \pi^h - \gamma \sum_{h,i} F_i \left[ \frac{\partial x_{i,j}^h}{\partial q_j} - \frac{\partial x_{i,j}^h}{\partial I} x_i^h \right] \pi^h = 0$$

$$\sum \lambda^h \frac{\partial V^h}{\partial I} \pi^h - \gamma \sum_{h,i} F_i \frac{\partial x_i^h}{\partial I} \pi^h = 0$$

- notation:
  - population average: $\mathbb{E}_h [\cdot] = \sum_h [\cdot] \pi^h$
  - adjusted pareto weight: $\beta^h \equiv \frac{\lambda^h \frac{\partial V^h}{\partial I}}{\gamma}$
we arrive at the condition

$$\mathbb{E}_h \left[ \sum_l t_l \frac{\partial x^c_{j}}{\partial q_l} \right] = X_j \mathbb{E}_h \left[ \frac{x^h_j}{X_j} \left( -1 + \beta^h + \sum_l t_l \frac{\partial x^h_l}{\partial I} \right) \right]$$

• Note that if we have homothetic and separable preferences then

$$\frac{x^h_j}{X_j}$$

is independent of $j$. So from here we can see a uniform tax result.

• if we have a lump sum then:

$$\mathbb{E}_h \left[ -1 + \beta^h + \sum_l t_l \frac{\partial x^h_l}{\partial I} \right] = 0$$

so we can write

$$\mathbb{E}_h \left[ \sum_l t_l \frac{\partial x^c_{j}}{\partial q_l} \right] = X_j \text{Cov}_h \left[ \frac{x^h_j}{X_j}, \beta^h \right]$$

where $\hat{\beta}^h = \beta^h + \sum_l t_l \frac{\partial x^h_l}{\partial I}$.

• We get two intuitive cases:
  - $\hat{\beta}^h$ is constant;
  - $\frac{x^h_j}{X_j}$ is independent of $j$. Then back to regular case.

• Q: Pareto inefficiency?

• A: If $\#agents < \#goods$ maybe cannot find $\beta^h$ that solve these equations

• Suppose utility is

$$U^i(G(x_1, \ldots, x_{N_1}), H(x_{N_1+1, \ldots, x_N}))$$

and $G, H$ are h.o.d. 1

• Result: tax uniformly within each group.

• Proof: treat goods $(x_1, x_2, \ldots, x_{N_1})$ and $(x_{N_1}, x_2, \ldots, x_N)$ as inputs into production of $G$ and $H$. 
3 Mixed Taxation: Atkinson-Stiglitz

- Notation:
  \( x \in \mathbb{R}^m \) consumption goods
  \( Y \in \mathbb{R} \) labor (in efficiency units)
  \( B \) budget set

- Given \( B \) consumers solve:
  \[
  (x^i, Y^i) \in \arg \max_{(x,Y) \in B} U^i(x, Y)
  \]

- Technology (linear)
  \[
  \sum_{i,j} p_j x^i_j \pi^i \leq \sum_i Y^i
  \]

- Feasibility. previous 2 conditions hold.

- if \( B^i \) allowed to be dependent on \( i \) then we can get the first best (Welfare theorem)

- ...but here \( B \) is independent of \( i \) so we are in the second best

- Assume:
  \[
  u^i(x, Y) = U^i(G(x), Y)
  \]

- Result: uniform taxation is efficient (Atkinson-Stiglitz).
  \[
  B_{AS} \equiv \{(x, Y) | p \cdot x \leq Y - T(Y)\}
  \]

  Indeed, anything else is Pareto inefficient!

- Exercise to get to result...
  1. start from \( B_0 \) that uses commodity taxes
  2. create new \( B \) that is “better”

  Here “better”: save resources and same utility. (Why better?)

- really can start from any arbitrary \( B_0 \)

- Note: “two stage” budgeting (given any \( B \))...
  Define:
  \[
  b = \{(g, Y) | \exists x \text{ s.t. } g = G(x) \text{ and } (x, Y) \in B\}
  \]
then agents solve (outer stage):

\[
\arg\max_{g,Y\in b} U^i(g, Y)
\]

- Idea: given \( B_0 \) we have some \( b_0 \). We change \( B_1 \) but keep implied \( b_1 = b_0 \). Then we get the same choices of \( Y^i \) and the same utility for each agent. Good choice:

\[
B_1 = B_{AS} \equiv \{ (x, Y) | \exists g \text{ s.t. } p \cdot x \leq e^G(g, p) \text{ and } (g, Y) \in b_0 \}
\]

where \( e^G(g, p) \equiv \min_x p \cdot x \text{ s.t. } g = G(x) \), is the expenditure function for \( G \).

- Equivalently if we define

\[
\hat{b} \equiv \{ (y, Y) | \exists g \text{ s.t. } y = e^G(g, p) \text{ and } (g, Y) \in b \}
\]

then

\[
B_{AS} \equiv \{ (x, Y) | p \cdot x \leq y \text{ and } (y, Y) \in \hat{b} \}
\]

which has an obvious income tax interpretation.

- This will save resources as long as \( x \) choices change. Why?

4 Pigouian Taxation

- now assume
  - single agent
  - lump sum taxation
  - but externalities

- utility

\[
u(x, \bar{x})
\]

concave in both \( x \) and \( \bar{x} \)

- technology

\[
F(x + g) = 0
\]

- in equilibrium

\[
\bar{x} = x
\]
agent solves (takes $\bar{x}$ as given)

$$\max_{x} u(x, \bar{x}) \quad q \cdot x = I$$

$$\Rightarrow u_{x}(x^e, x^e) = \lambda q$$

$$\Rightarrow \frac{q_i}{q_j} = \frac{u_{x_i}(x^e, x^e)}{u_{x_j}(x^e, x^e)}$$

Social optimum

$$\max_{x} u(x, x) \quad F(x + g) = 0$$

$$\Rightarrow u_{x}(x^*, x^*) + u_{\bar{x}}(x^*, x^*) = \gamma F_{x}(x^* + g)$$

$$\Rightarrow \frac{p_i}{p_j} = \frac{F_{x_i}}{F_{x_j}} = \frac{u_{x_i}(x^*, x^*) + u_{\bar{x}_i}(x^*, x^*)}{u_{x_j}(x^*, x^*) + u_{\bar{x}_j}(x^*, x^*)}$$

To make

$$x^e = x^*$$

a necessary condition is that both conditions hold, implying

$$\frac{p_i}{q_i} / \frac{p_j}{q_j} = \frac{1 + \frac{u_{x_i}(x^*, x^*)}{u_{x_j}(x^*, x^*)}}{1 + \frac{u_{\bar{x}_i}(x^*, x^*)}{u_{\bar{x}_j}(x^*, x^*)}}$$

Theorem: if $p$ and $q$ to satisfy this equation, then there exists an income $I$ (i.e. lump sum tax/transfer) so that the agent chooses $x = x^*$.

Proof: (sketch) Use Lagrangian sufficiency theorem.

5 Application to Intertemporal Taxation

- neoclassical growth model

- simplifying assumptions
  - single agent first
  - no uncertainty
• technology

\[ c_t + g_t + k_{t+1} \leq F(k_t, L_t) + (1 - \delta)k_t \]

where \( F \) is CRS

• preferences

\[ \sum_{t=0}^{\infty} \beta^t u(c_t, L_t) \]

• budget constraints

  - agents

\[ c_t + k_{t+1} + q_{t,t+1}B_{t+1} \leq (1 - \tau_t)w_tL_t + R_tk_t + (1 - \kappa_t^B)B_t \]

where

\[ R_t = 1 + \kappa_t(r_t - \delta) \]

we also need some no-ponzi conditions

\[ q_{0,t} = q_{0,1}q_{1,2}\cdots q_{t-1,t} \]

\[ \lim_{T \to \infty} q_{0,T}B_T \geq 0 \]

  - government:

\[ g_t + B_t \leq \tau_t w_tL_t + \kappa_t r_tk_t + q_{t,t+1}B_{t+1} \]

• without loss of generality:

\[ \kappa_t^B = 0 \quad t = 1, 2, \ldots \]

• Firms:

\[ \max_{K_t, L_t} \left\{ F(K_t, L_t) - w_tL_t - r_tK_t \right\} \]

necessary and sufficient conditions

\[ F_L(K_t, L_t) = w_t \]

\[ F_K(K_t, L_t) = r_t \]

• Definition of an equilibrium:

  - agents maximize given prices and taxes
- firms maximize
- government budget constraint satisfied
- market clears: goods, capital and bonds

• adding up both budget constraints gives

\[ g_t + c_t + k_{t+1} \leq w_t L_t + (1 + r_t - \delta)k_t = F(k_t, L_t) + (1 - \delta)k_t \]

which is just the resource constraint

• solving \( B_t \) forward

\[ \sum_{t=0}^{\infty} q_{0,t} (c_t - (1 - \tau_t)w_t L_t - R_t k_t + k_{t+1}) \leq (1 - \kappa_0^B)B_0 \]

unless

\[ q_{t+1}R_{t+1} = \frac{q_{0,t+1}}{q_{0,t}} R_{t+1} = 1 \quad t = 0, 1, \ldots \]

there is an arbitrage

• cancelling:

\[ \sum_{t=0}^{\infty} q_{0,t} (c_t - (1 - \tau_t)w_t L_t) \leq R_0 k_0 + (1 - \kappa_0^B)B_0 \]

• now we can just apply the primal approach

• implementability condition:

\[ \sum_{t=0}^{\infty} \beta^t (u_{c,t} c_t + u_{L,t} L_t) = u_{c,0}(R_0 k_0 + (1 - \kappa_0^B)B_0) \]

• Lagrangian

\[ L \equiv \sum_{t=0}^{\infty} \beta^t W(c_t, L_t; \mu) - \mu u_{c,0}(R_0 k_0 + (1 - \kappa_0^B)B_0) \]

where

\[ W(c, L; \mu) \equiv u(c, L) + \mu (u_c(c, L)c + u_L(c, L)L) \]

• optimality conditions obtained from

\[ \max L \quad \text{s.t. resource constraint} \]
• first order conditions:

\[- \frac{W_L(c_t, L_t; \mu)}{W_c(c_t, L_t; \mu)} = F_L(K_t, L_t)\]

\[W_c(c_t, L_t; \mu) = \beta R^*_{t+1} W_c(c_{t+1}, L_{t+1}; \mu)\]

where \( R^*_{t+1} \equiv F_k(k_{t+1}, L_{t+1}) + 1 - \delta \) is the social rate of return

• for agent

\[w_t(1 - \tau_t) = -\frac{u_L(c_t, L_t)}{u_c(c_t, L_t)}\]

\[u_c(c_t, L_t) = \beta R_t u_c(c_{t+1}, L_{t+1})\]

• implications

\[1 - \tau_t = \frac{u_L(c_t, L_t)}{u_c(c_t, L_t)} \frac{W_c(c_t, L_t; \mu)}{W_L(c_t, L_t; \mu)}\]

\[\frac{R_{t+1}}{R^*_{t+1}} = \frac{u_c(c_t, L_t)}{u_c(c_{t+1}, L_{t+1})} \frac{W_c(c_{t+1}, L_{t+1}; \mu)}{W_c(c_t, L_t; \mu)}\]

• results:

  – a form of labor tax smoothing:

  * the entire sequence of \( g_t \) has an impact on the tax through \( \mu \)
  * no special role for current \( g_t \), conditional on current allocation
  * clearer in special cases: if

\[u(c, L) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{\alpha L^{\gamma}}{\gamma}\]

with \( \sigma > 0 \) and \( \gamma > 1 \) then

\[\tau_t = \bar{\tau}\]

  – at a steady state the tax on capital is zero (Chamley-Judd):

\[c_t \rightarrow \bar{c} \quad L_t \rightarrow \bar{L}\]

\[\Rightarrow \frac{R_{t+1}}{R^*_{t+1}} \rightarrow 1\]

  – initial tax on capital and bonds:

  * equivalent to a lump sum tax
  * optimal to expropriate
• if upper bound on tax rates, then they will binding

• last result leads to time inconsistency:
  – plan to...
    * tax initial capital highly
    * tax future capital at zero
  – will plan be carried out? can we commit to it?
    * if not, and reoptimize once and for all then raise capital again
    * if reoptimize all the time (discretion): expect high taxes, which lowers welfare

• with heterogeneous agents
  – allow a lump sum (poll) tax
  – first two results hold: tax smoothing and Chamley-Judd
  – the last conclusion less clear:
    * Pareto analysis
      * depends on distribution of assets and redistributive intent
  – even if capital levy is optimal, it may be bounded, and correct intuition is not based on a lump sum tax
  – time inconsistency also more subtle: in general not time consistent, but depends on evolution of wealth