Recitation 1: Pigouvian Taxation

Todd Lensman

February 4, 2022

Recitation Plan: Discuss and solve examples of the general Pigouvian taxation model

1 General Model with a Representative Consumer

Consumption. The economy has a continuum of identical consumers. Each consumer has preferences defined over her own net consumption vector $x \in X$ and the average net consumption vector in the population $\bar{x} \in X$, where $X \subseteq \mathbb{R}^n$ is assumed convex with a non-empty interior. These preferences are represented by a utility function $u(x, \bar{x})$, assumed differentiable in all arguments and concave and locally non-satiated in x for any $\bar{x}$. Since $\bar{x}$ directly enters u, we have the possibility of consumption externalities – one consumer's choice can directly affect another's utility.

Given consumer prices $q \in \mathbb{R}^n$, a lump-sum tax $T \in \mathbb{R}$, and the average consumption choice $\bar{x}$, the representative consumer solves

$$\max_{x \in X} u(x, \bar{x}) \quad \text{subject to} \quad q \cdot x + T \leq 0. \tag{1.1}$$

Throughout, we assume that the solution occurs at an interior point of X, so the budget constraint is the only active constraint.

Example 1.1. Consider an economy that only has two "goods," consumption c and labor l. The representative consumer has endowments $\tilde{c}$ and $\tilde{l}$ of each, and we naively define the consumer's preferences by the uility function $v(c, l)$, where v is concave, strictly increasing in c, and strictly decreasing in l (note that there are no externalities). Suppose we also impose that final consumption must be non-negative, while labor supply must be non-negative and weakly smaller than the consumer's labor endowment. The consumer's problem is then

$$\max_{c \geq 0, \tilde{l} \geq l \geq 0} v(c, l) \quad \text{subject to} \quad q_c(c - \tilde{c}) + T \leq q_l l. \tag{1.2}$$

To cast this specification of consumption in terms of the general model, we recall that x is interpreted as the vector of transacted quantities in the market. That is, $x_c = c - \tilde{c}$ and $x_l = -l$, where x_l is negative to maintain the convention that prices are non-negative. Then the consumption set is $X := \left\{ x \mid x_c \geq -\tilde{c}, 0 \geq x_l \geq -\tilde{l} \right\}$, and the consumer's utility function over x is $u(x) := v(x_c + \tilde{c}, -x_l)$.

Production. The economy's production technology is described by the transformation function $F : \mathbb{R}^n \to \mathbb{R}$, where a net output vector is feasible if and only if $F(y) \leq 0$. We assume that F is differentiable and homogeneous of degree one, so that the production technology has constant returns to scale. With price-taking firms, this has the useful implication that "market structure" does not matter: Any number of firms may be operating with the same technology, or the firms may have heterogeneous technologies provided that each satisfies constant returns to scale. In both cases, we can describe the aggregate production technology for the economy using a transformation function F.

Given the transformation function F and producer prices p, the representative competitive firm maximizes profits over production vectors y:

$$\max_{y \in \mathbb{R}^n} p \cdot y \quad \text{subject to} \quad F(y) \leq 0. \tag{1.3}$$

Example 1.2. Consider an economy with two produced goods $\{1, 2\}$ and one labor good. We naively describe the production technology for the economy using production functions:

$$\tilde{y}_1 = A_1 \tilde{y}_2^{\alpha} \tilde{L}_1^{1-\alpha} \quad \text{and} \quad \tilde{y}_2 = A_2 \tilde{L}_2. \tag{1.4}$$

That is, labor is used to produce goods 1 and 2, and good 2 is additionally used to produce good 1. What is the induced transformation function $F(y_1, y_2, -L)$?[1] We can find a candidate by attempting to maximize the net output of one of the goods (say y_1) while respecting the constraints imposed by the production functions described above:

$$F(y_1, y_2, -L) \quad := \quad y_1 - \max_{\tilde{L}_1, \tilde{L}_2, \tilde{y}_2 \geq 0} A_1 \tilde{y}_2^{\alpha} \tilde{L}_1^{1-\alpha} \tag{1.5}$$

$$\text{subject to} \tag{1.6}$$

$$y_2 = A_2 \tilde{L}_2 - \tilde{y}_2 \tag{1.7}$$

$$L = \tilde{L}_1 + \tilde{L}_2. \tag{1.8}$$

[1] But note that F is not uniquely defined: Given any F, $\varphi \circ F$ for any strictly increasing function φ with $\varphi(0) = 0$ is a transformation function that describes the same production technology.

To interpret, the maximization problem requires that we maximize the net output of good 1, subject to producing net outputs $(y_2, -L)$ of the remaining goods.[2] $F(y_1, y_2, -L)$ is then the difference between the prescribed net output y_1 and the maximal net output of good 1. It is easy to verify that F is homogeneous of degree one and that $F(y_1, y_2, -L) \leq 0$ characterizes the set of production vectors that are feasible given the production functions (1.4) and free disposal.

Government. The government uses commodity taxes and a lump-sum tax to finance a vector of public spending $g \in \mathbb{R}^n$ and potentially correct market failures due to externalities. The government's implied budget constraint is

$$(q-p) \cdot x + T = p \cdot g. \tag{1.9}$$

Equilibrium. An equilibrium is this economy is essentially a Walrasian competitive equilibrium with taxes: a tuple (x, q, p, T) such that

(i) the government's budget constraint is satisfied;

(ii) the representative consumer chooses consumption vector x given q, T, and $\bar{x} = x$;

(iii) the representative firm chooses production vector y given p; and

(iv) all markets clear, $x + g = y$.

Pigouvian Tax Formula. The key optimality condition for relative prices in this environment is

$$\frac{p_i/q_i}{p_j/q_j} = \frac{1 + \frac{u_{\bar{x}_i}(x_*, x_*)}{u_{x_i}(x_*, x_*)}}{1 + \frac{u_{\bar{x}_j}(x_*, x_*)}{u_{x_j}(x_*, x_*)}}. \tag{1.10}$$

This formula indicates that the marginal rate of substitution between i and j, q_i/q_j, should be distorted downward from the marginal technical rate of substitution p_i/p_j whenever the normalized marginal externality from good i, $u_{\bar{x}_i}/u_{x_i}$, is large relative to the normalized marginal externality from good j, $u_{\bar{x}_j}/u_{x_j}$.

[2] The similarity with a standard Pareto optimality problem is no accident! Whether we are interested in characterizing the production possibilities frontier (as we are here) or the Pareto frontier in a given economy, the goal is to optimize one component of a vector subject to constraints on the remaining components.

2 Examples

Example 2.1. Suppose the economy has two goods, consumption c and labor l, and suppose that the representative consumer has preferences given by the utility function

$$v(c,l,\bar{c}) = c - \frac{l^{1+\eta}}{1+\eta} - \alpha\bar{c}. \tag{2.1}$$

Here aggregate consumption $\bar{c}$ imposes an additively separable externality on the representative consumer. The production technology is linear: $F(c,-l) = c - Al$.

To solve for the optimal corrective taxes, we begin by solving the competitive equilibrium given policy variables (q,p,T). The consumer's problem is

$$\max_{c,l\geq 0} c - \frac{l^{1+\eta}}{1+\eta} - \alpha\bar{c} \quad \text{subject to} \quad q_c c - q_l l + T \leq 0. \tag{2.2}$$

The first-order conditions imply

$$l = \left(\frac{q_l}{q_c}\right)^{\frac{1}{\eta}} \quad \text{and} \quad c = \left(\frac{q_l}{q_c}\right)^{\frac{1+\eta}{\eta}} - \frac{T}{q_c}. \tag{2.3}$$

Note that since the externality is additively separable in the consumer's utility function, it has no direct impact on the consumer's equilibrium consumption or labor supply decisions. The representative firm's problem is

$$\max_{c,l} p_c c - p_l l \quad \text{subject to} \quad c \leq Al. \tag{2.4}$$

The first-order condition implies $p_c = p_l/A$. Finally, recall the government's budget constraint:

$$(q_c - p_c)c - (q_l - p_l)l + T = p_c c^G + p_l l^G. \tag{2.5}$$

This equation can be solved simultaneously with the consumer's consumption choice to express the tax T as a function of prices (q,p) and exogenous government consumption $\left(c^G, l^G\right)$.

To determine optimal values for the policy variables (q,p,T), we solve the government's welfare maximization problem. In particular, the government optimizes over consumption c and labor l, internalizing the consumption externality and subject to the production technology:

$$\max_{c,l} (1-\alpha)c - \frac{l^{1+\eta}}{1+\eta} \quad \text{subject to} \quad c + c^g \leq A(l - l^g). \tag{2.6}$$

The first-order conditions imply

$$l^* = [(1-\alpha)A]^{\frac{1}{\eta}} \quad \text{and} \quad c^* = A\Big([(1-\alpha)A]^{\frac{1}{\eta}} - l^g\Big) - c^g. \tag{2.7}$$

To implement this allocation in equilibrium, we must set

$$\frac{q_l}{q_c} = (1-\alpha)A \quad \text{and} \quad \frac{p_l}{p_c} = A, \tag{2.8}$$

with the lump-sum tax T chosen to satisfy the government's budget constraint given allocation (c^*, l^*), chosen prices (q, c), and exogenous government consumption $\left(c^G, l^G\right)$. Note in particular that we arrive at one relation between relative prices that determines the optimal tax distortion:

$$\frac{p_c/q_c}{p_l/q_l} = 1 - \alpha = \frac{1 + \frac{v_{\bar{c}}(c^*,l^*,c^*)}{v_c(c^*,l^*,c^*)}}{1 + \frac{v_{\bar{l}}(c^*,l^*,c^*)}{v_l(c^*,l^*,c^*)}}. \tag{2.9}$$

The last equality emphasizes the connection with the general Pigouvian tax formula.[3] We can implement this relation (and the first-best allocation) using a variety of different tax instruments. For example, taxing consumption suffices:

$$q_c = (1+\tau_c)\,p_c \quad \text{and} \quad q_l = p_l, \tag{2.10}$$

where $1 + \tau_c = 1/(1-\alpha)$. This tax increases the consumer's price of consumption and distorts downward equilibrium consumption, aligning the private and social marginal benefits of consumption. These two equations, along with the producer optimality condition $p_c = p_l/A$ and a standard Walrasian price normalization, determine equilibrium consumer and producer prices. The lump-sum tax T can then be computed from the government's budget constraint. Alternatively, we can tax labor supply:

$$q_c = p_c \quad \text{and} \quad q_l = (1+\tau_l)\,p_l, \tag{2.11}$$

were $1 + \tau_l = 1 - \alpha$. Again, this tax has the effect of distorting downward equilibrium consumption.

[3]And note in this case how much easier it is to derive the optimal price distortions from the general formula than by solving for the equilibrium and the social optimum! This only holds because there are no externalities associated with labor, and the marginal utility of consumption and the marginal consumption externality are both constant.

Example 2.2. In this example, we consider conditions under which the general equilibrium analysis above can be reduced to the partial equilibrium graphical analysis often seen in undergraduate classes. We begin with the general model, and we suppose that the representative consumer's utility function has the additively separable, quasilinear representation

$$u(x,\bar{x}) = x_1 + \sum_{i=2}^{n} u^i(x_i,\bar{x}_i). \tag{2.12}$$

The first-order conditions to the consumer's problem are then

$$x_1 = \sum_{i=2}^{n} \frac{q_i}{q_1} x_i - \frac{T}{q_1}, \tag{2.13}$$

$$u^i_{x_i}(x_i,\bar{x}_i) = q_i \quad i \in \{2,\ldots,n\}. \tag{2.14}$$

Now $\bar{x}_i = x_i$ in equilibrium, so we find that the equilibrium inverse demand function $x_i \mapsto u^i_{x_i}(x_i,x_i)$ is independent of the quantities in markets $j \neq i$ (because of additive separability) and the consumer's wealth (because of quasilinearity). Similarly, the "social marginal benefit" function $x_i \mapsto u^i_{x_i}(x_i,x_i) + u^i_{\bar{x}_i}(x_i,x_i)$ is also independent of the quantities in markets $j \neq i$. Suppose also that the aggregate transformation function takes the separable form

$$F(y) = \tag{2.15}$$

Recitation 2: Diamond-Mirrlees I

Todd Lensman

February 11, 2022

Recitation Plan: Review the Diamond-Mirrlees production efficiency result and applications

1 General Model with a Representative Consumer

Consumption. The economy has a continuum of identical consumers. Each consumer has preferences defined over her own net consumption vector $x \in X$, where $X \subseteq \mathbb{R}^n$ is assumed convex with a non-empty interior. These preferences are represented by a utility function $u(x)$, assumed differentiable, concave, and locally non-satiated.

Given consumer prices $q \in \mathbb{R}^n$ and a lump-sum tax $T \in \mathbb{R}$, the representative consumer solves

$$\max_{x \in X} u(x) \quad \text{subject to} \quad q \cdot x + T \leq 0. \tag{1.1}$$

Throughout, we assume that the solution occurs at an interior point of X, so the budget constraint is the only active constraint.

Private Production. The private sector's production technology is described by the transformation function $F : \mathbb{R}^n \to \mathbb{R}$, where a net output vector is feasible if and only if $F(y) \leq 0$. We assume that F is differentiable and homogeneous of degree one, so that the production technology has constant returns to scale.

Given the transformation function F and producer prices p, the representative competitive firm maximizes profits over production vectors y:

$$\max_{y \in \mathbb{R}^n} p \cdot y \quad \text{subject to} \quad F(y) \leq 0. \tag{1.2}$$

Government. The government uses commodity taxes and a lump-sum tax to finance a vector of public spending $g \in \mathbb{R}^n$, and it can also engage in public production of commodities. Let $z \in \mathbb{R}^n$ denote the vector of net government production, inclusive of spending g, and let $G : \mathbb{R}^n \to \mathbb{R}$

denote the government's transformation function. The government's production constraint is $G(z) \leq 0$, and its implied budget constraint is

$$(q - p) \cdot x + T + p \cdot z = 0. \tag{1.3}$$

Equilibrium. An equilibrium is this economy is essentially a Walrasian competitive equilibrium with taxes: a tuple (x, z, q, p, T) such that

(i) the government's budget constraint is satisfied;

(ii) the government's production constraint is satisfied;

(iii) the representative consumer chooses consumption vector x given q and T;

(iv) the representative firm chooses production vector y given p; and

(v) all markets clear, $x = y + z$.

2 Production Efficiency + Applications

2.1 The Result

Toward the production efficiency result of Diamond and Mirrlees (1971a), we restrict the government's policy tools by setting $T = 0$. Otherwise, the government can optimally finance any (net) government spending by levying a lump-sum tax on the representative consumer and leaving all relative prices undistorted – this is a direct implication of the First Welfare Theorem. In this case, aggregate production efficiency (i.e., $F(y) = 0$ and $G(z) = 0$) is necessarily satisfied at the optimum.

With $T = 0$, the welfarist government's problem is as follows:

$$\max_{z,q,p} u(x(q)) \quad \text{subject to} \quad G(z) \leq 0, \tag{2.1}$$

$$x(q) = y(p) + z. \tag{2.2}$$

Here $x(q)$ denotes the solution to the consumer's problem given q and $T = 0$, and $y(p)$ denotes the solution to the firm's problem given p. Note that since the consumer's budget constraint is necessarily satisfied with equality and the firm makes zero profits, the market-clearing conditions in the government's problem immediately imply that the government's budget constraint is satisfied with $T = 0$.

To simplify this problem, we observe that since the firm's problem is concave, the solution is characterized by the first-order conditions

$$\frac{p_i}{p_1} = \frac{F_i}{F_1} \quad i \in \{1, \ldots, n\}. \tag{2.3}$$

Equilibrium with $p \neq 0$ also implies that the firm must be productively efficient, $F(y) = 0$. As we have seen in lecture, a key implication of this fact is that the government can obtain any feasible and productively efficient net output vector y from the private sector with an appropriate choice of producer prices p. We can then equivalently write the government's problem without producer prices p but with a government choice of private sector net output y:

$$\max_{y,z,q} u(x(q)) \quad \text{subject to} \quad F(y) \leq 0, \tag{2.4}$$

$$G(z) \leq 0, \tag{2.5}$$

$$x(q) = y + z. \tag{2.6}$$

Note that I have already incorporated a relaxation by allowing the government to choose a productively inefficient private net output vector (with $F(y) < 0$). The following result demonstrates that this relaxation does not change the solution to the government's problem:

Theorem 2.1 (Productive Efficiency)**.** Let (y, z, q) be a solution to the government's problem with $x_k(q) \neq 0$. Then

$$F(y) = 0, \qquad G(z) = 0, \qquad \frac{F_i}{F_1} = \frac{G_i}{G_1} \quad i \in \{1, \ldots, n\}. \tag{2.7}$$

Proof. Let λ^F and λ^G be the multipliers on the private and public production constraints, and let γ_i be the multiplier on the market-clearing constraint for commodity i. Then the first-order conditions with respect to y_i and z_i are

$$\lambda^F F_i = \gamma_i = \lambda^G G_i. \tag{2.8}$$

Provided that the multipliers λ^F and λ^G are non-zero, and that the transformation functions F and G are strictly monotone in some commodity, we can conclude. To establish these facts, note that the first-order condition with respect to q_k can be written

$$\alpha x_k(q) = \sum_{i=1}^{n} \gamma_i \frac{\partial x_i}{\partial q_k}. \tag{2.9}$$

where $\alpha > 0$ denotes the consumer's marginal utility of wealth. Hence there exists some commodity, say 1, such that $\gamma_1 \neq 0$. This implies $\lambda^F, \lambda^G, F_1, G_1 \neq 0$. ■

Remark. In the version of the model with heterogeneous consumers, a nearly identical argument demonstrates that productive efficiency must hold at the optimum. Except for technical conditions that ensure the Walrasian equilibrium is well-behaved, the only modification we must make to our assumptions is that there exists some commodity k such that $x_k^h(q) \geq 0$ for all households h or $x_k^h(q) \leq 0$ for all households h, with a strict inequality for some h. This implies that the multiplier γ_k is again nonzero. Try re-doing the proof with heterogeneous consumers to see how this works, or check out Section IV of Diamond & Mirrlees (1971a) for full technical details. Alternatively, with heterogeneous consumers, the argument becomes even easier if we allow a uniform lump-sum tax/subsidy (uniformity implies that the optimum may still be second-best).

2.2 Applications

Management of Public Production

What objective should be assigned to a publicly-owned firm when the government seeks to maximize welfare over commodity taxes and the public production vector? The production efficiency theorem immediately implies that public and private marginal technical rates of substitution should be equated, so the publicly-owned firm should act exactly like a privately-owned firm (but with technology G instead of F).

Intermediate Taxation

Should we ever tax firm-to-firm transactions? Not according to the production efficiency result! To see this, we can reinterpret F and G as the production technologies available in two different private sectors (e.g., manufacturing and services). Suppose, as is necessary for intermediate good taxation, that the government can assign different price vectors p^F and p^G to sectors F and G, respectively. Then by the same argument as above, the government can use prices p^F and p^G to obtain any productively efficient net output vectors y and z from sectors F and G. Relaxing the requirement of within-sector productive efficiency, we obtain the same optimization problem for the government that we analyzed in the proof of the production efficiency result. The conclusion: each sector should be internally productively efficient, and we should equate marginal rates of technical substitution across sectors. To see how these facts imply aggregate productive efficiency, we can analyze the aggregate production technology and characterize its frontier. In particular, let $a = y + z$ denote aggregate net output, and note that the aggregate technology is described by the transformation function

$$
\begin{aligned}
A(a) \quad &:= \quad a_1 - \max_{y,z} y_1 + z_1 && (2.10)\\
&\qquad \text{subject to} && (2.11)\\
&\qquad y_i + z_i \geq a_i \quad i \in \{2, \ldots, n\}, && (2.12)\\
&\qquad F(y) \leq 0, && (2.13)\\
&\qquad G(z) \leq 0. && (2.14)
\end{aligned}
$$

The optimization problem is convex, so the solution is characterized by the first-order conditions. In particular, we must have

$$
\begin{aligned}
1 &= \lambda^F F_1 = \lambda^G G_1, && (2.15)\\
\gamma_i &= \lambda^F F_i = \lambda^G G_i \quad i \in \{2, \ldots, n\}. && (2.16)
\end{aligned}
$$

Under weak monotonicity conditions on the production technologies,[1] the Lagrange multipliers λ^F and λ^G are non-zero, and the marginal technical rates of substitution are equated across sectors. The aggregate production frontier is then characterized by the conditions

$$
F(y) = 0, \qquad G(z) = 0, \qquad \frac{F_i}{F_k} = \frac{G_i}{G_k} \quad i \in \{1, \ldots, n\}. \tag{2.17}
$$

But these are precisely the conditions implied by the production efficiency theorem! Thus the production efficiency theorem implies productive efficiency in each private sector and equality of marginal technical rates of substitution across sectors (i.e., no intermediate taxation across sectors), which in turn are equivalent to aggregate production efficiency.

Trade

Should a domestic welfare-maximizing government ever tax or subsidize imports or exports? The production efficiency theorem again says no, assuming the government acts as a price-taking firm on the world market. To see this, suppose that commodities 1 and 2 are tradable,

[1] These essentially require that "you can't get something for nothing."

and suppose the home country faces international prices p_1^I and p_2^I. From the perspective of the home country, trade amounts to a private production technology that allows the exchange of commodity 1 for commodity 2 at marginal technical rate of substitution p_1^I/p_2^I. Formally, we can reinterpret technology G as that of a private importer-exporter:

$$G(z_1, z_2) = p_1^I z_1 + p_2^I z_2. \tag{2.18}$$

Proceeding according to the proof of the production efficiency theorem, we can conclude that the solution must feature

$$\frac{G_1}{G_2} = \frac{p_1^I}{p_2^I} = \frac{F_1}{F_2}. \tag{2.19}$$

In particular, the government should not distort marginal technical rates of substitution away from those that prevail at international prices. For example, suppose that commodity 1 is produced by domestic firms using commodity 2 as an input. The government should not subsidize domestic output of commodity 1, because the last unit of commodity 2 used in domestic production could be more efficiently used by exchanging it for commodity 1 on the world market. The "return" in units of commodity 1 from international trade is more favorable because domestic firms over-produce commodity 1 due to the subsidy, leading to diminished marginal returns in production.

Recitation 3: Diamond-Mirrlees II + Pigou

Todd Lensman

February 18, 2022

Recitation Plan: Integrate externalities into the Diamond-Mirrlees II optimal tax formulas and derive the Sandmo (1975) additivity result

1 General Model

Consumption. The economy has a continuum of heterogeneous consumers, where each consumer has a type h that belongs to a finite set H. Let π^h denote the proportion of consumers of type h. Each consumer has preferences defined over her own net consumption vector $x^h \in X$ and the average net consumption vector in the population $\bar{x} \in X$, where $X \subseteq \mathbb{R}^n$ is assumed convex with a non-empty interior. These preferences are represented by a utility function $u^h\left(x^h, \bar{x}\right)$, assumed differentiable in all arguments and concave and locally non-satiated in x^h for any $\bar{x}$. Since $\bar{x}$ directly enters u, we have the possibility of consumption externalities – one consumer's choice can directly affect another's utility.

Given consumer prices $q \in \mathbb{R}^n$, a lump-sum tax $T^h \in \mathbb{R}$, and the average consumption choice $\bar{x}$, a type h consumer solves

$$\max_{x^h \in X} u^h\left(x^h, \bar{x}\right) \quad \text{subject to} \quad q \cdot x^h + T^h \leq 0.$$

Throughout, we assume that the solution occurs at an interior point of X, so the budget constraint is the only active constraint.

Production. The production technology is described by the transformation function $F : \mathbb{R}^n \to \mathbb{R}$, where a net output vector is feasible if and only if $F(y) \leq 0$. We assume that F is differentiable and homogeneous of degree one, so that the production technology has constant returns to scale.

Given the transformation function F and producer prices p, the representative competitive firm maximizes profits over production vectors y:

$$\max_{y \in \mathbb{R}^n} p \cdot y \quad \text{subject to} \quad F(y) \leq 0.$$

Government. The government uses commodity taxes and a uniform lump-sum tax to finance a vector of public spending $g \in \mathbb{R}^n$ and potentially correct market failures due to externalities. The government's implied budget constraint is

$$(q - p) \cdot \bar{x} + T = p \cdot g.$$

Equilibrium. An equilibrium is this economy is essentially a Walrasian competitive equilibrium with taxes: a tuple $\left(\left(x^h\right)_{h \in H}, z, q, p, T\right)$ such that

(i) the government's budget constraint is satisfied;

(ii) each type h consumer chooses consumption vector x^h given q and T;

(iii) the representative firm chooses production vector y given p; and

(iv) all markets clear, $\sum_h x^h \pi^h + g = y$.

2 Optimal Tax Formulas

Below we derive optimal tax formulas generalizing those of Diamond and Mirrlees (1971), and we use them to study how externalities affect the structure of optimal commodity taxes. Following the dual approach, let $V^h(q, I, \bar{x})$ denote the indirect utility function for type h consumers, and let $x^h(q, I, \bar{x})$ and $\hat{x}^h(q, u, \bar{x})$ denote the Marshallian and Hicksian demand functions. Making use of arguments from previous recitations to justify government control of production, we can write the government's second-best Pareto problem as follows:

$$\max_{q, I} \mathbb{E}_h\left[V^h(q, I, \bar{x}) \lambda^h\right] \quad \text{subject to} \quad F(\bar{x} + g) \leq 0,$$
$$\bar{x} = \mathbb{E}_h\left[x^h(q, I, \bar{x})\right].$$

Here $\mathbb{E}_h$ denotes an expectation with respect to the distribution of types. This problem differs from the standard commodity taxation problem only by the inclusion of $\bar{x}$ in the indirect utility and demand functions. But note that $\bar{x}$ must satisfy its own fixed-point equation, so to derive correct optimal tax formulas we have to be careful in how we deal with this additional condition. One option would be to include $\bar{x}$ as a choice variable for the government and keep the fixed-point equation for $\bar{x}$ as a constraint. I follow a different strategy: Keeping with the spirit of the dual approach, I appeal to the Implicit Function Theorem and let $\bar{x}(q,I)$ denote the solution to the fixed-point equation, which under appropriate conditions is unique and continuously differentiable in a neighborhood of the solution to the government's problem. The government's problem can then be written

$$\max_{q,I} \sum_{h\in H} V^h(q,I,\bar{x}(q,I))\lambda^h \pi^h \quad \text{subject to} \quad F(\bar{x}(q,I)+g) \leq 0.$$

Letting $\kappa > 0$ denote the multiplier on the feasibility constraint, the first-order condition with respect to q_i is

$$0 = \mathbb{E}_h\left[\lambda^h \frac{\partial V^h}{\partial q_i} + \lambda^h \frac{\partial V^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right] - \kappa \sum_{j=1}^{n} \frac{\partial F}{\partial y_j}\mathbb{E}_h\left[\frac{\partial x_j^h}{\partial q_i} + \frac{\partial x_j^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right].$$

Using Roy's Identity and the Slutsky Equation on the first and second terms, respectively, gives

$$0 = \frac{1}{\kappa}\mathbb{E}_h\left[-\lambda^h x_i^h \frac{\partial V^h}{\partial I} + \lambda^h \frac{\partial V^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right] - \sum_{j=1}^{n} \frac{\partial F}{\partial y_j}\mathbb{E}_h\left[\frac{\partial \hat{x}_j^h}{\partial q_i} - x_i^h\frac{\partial x_j^h}{\partial I} + \frac{\partial x_j^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right].$$

Substituting the producer price $p_j = \partial F/\partial y_j$ and using Slutsky symmetry and the homogeneity of degree zero identity $\sum_j q_j\left(\partial \hat{x}_i^h/\partial q_j\right) = 0$ on the second term, we can rearrange:

$$\begin{aligned}\sum_{j=1}^{n} t_j \mathbb{E}_h\left[\frac{\partial \hat{x}_j^h}{\partial q_i}\right] = \bar{x}_i \mathbb{E}_h&\left[\frac{x_i^h}{\bar{x}_i}\left(\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial I} - \sum_{j=1}^{n} p_j \frac{\partial x_j^h}{\partial I}\right)\right] \\ &- \mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i} - \sum_{j=1}^{n} p_j \frac{\partial x_j^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right]\end{aligned}$$

For one final manipulation, use the budget constraint identities $\sum_j q_j\left(\partial x_j^h/\partial I\right) = 1$ and $\sum_j q_j\left(\partial x_j^h/\partial \bar{x}\right) = 0$ to write

$$\sum_{j=1}^{n} t_j \mathbb{E}_h\left[\frac{\partial \hat{x}_j^h}{\partial q_i}\right] = \bar{x}_i \mathbb{E}_h\left[\frac{x_i^h}{\bar{x}_i}\left(\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial I} - 1 + \sum_{j=1}^{n} t_j \frac{\partial x_j^h}{\partial I}\right)\right] \tag{DM}$$

$$-\mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}+\sum_{j=1}^{n} t_j\frac{\partial x_j^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right].$$

This is precisely the optimal tax formula from Diamond and Mirrlees (1971), but with an additional term on the right side that accounts for the effect of the externalities on the optimal tax structure. For intuition (which also applies in the case with no externalities!), note that we could heuristically derive this equation by solving the following version of the government's problem:

$$\max_{q,I}\mathbb{E}_h\left[V^h\left(q,I,\bar{x}\left(q,I\right)\right)\lambda^h\right]\quad\text{subject to}\quad(q-p)\cdot\bar{x}-I=p\cdot g.$$

Here we maximize Pareto-weighted welfare over consumer prices and lump sum income, but subject to the government's budget constraint and *holding producer prices fixed*. Letting κ denote the multiplier on the budget constraint, the first-order condition with respect to q_i is

$$-\mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\left(-x_i^h\frac{\partial V^h}{\partial I}+\frac{\partial V^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right)\right]=\bar{x}_i+\sum_{j=1}^{n} t_j\mathbb{E}_h\left[\frac{\partial x_j^h}{\partial q_i}+\frac{\partial x_j^h}{\partial \bar{x}}\frac{d\bar{x}}{dq_i}\right]. \tag{DM'}$$

Note that the left side captures the direct effect of the price increase on indirect utilities, while the right side captures the effect on government revenues. This equation can be rearranged to give the Diamond-Mirrlees formula (DM), so the latter captures exactly the "direct effects and fiscal externalities" intuition discussed in class.

Continuing with formula (DM'), we can get a sense for how externalities affect optimal commodity taxes by considering a special case. Suppose the utility functions take the weakly separable form $u^h\left(x^h,\bar{x}\right)=\tilde{u}^h\left(w^h\left(x^h\right),\bar{x}\right)$. Then the Marshallian demand functions $x^h\left(q,I\right)$ do not depend on average consumption $\bar{x}$, and average consumption is no longer defined by a fixed-point equation:

$$\bar{x}\left(q,I\right)=\sum_{h\in H}x^h\left(q,I\right)\pi^h.$$

Applying these observations, the formula (DM') becomes

$$\mathbb{E}_h\left[\frac{\lambda^h}{\kappa}x_i^h\frac{\partial V^h}{\partial I}\right]-\sum_{j=1}^{n}\mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial \bar{x}_j}\right]\mathbb{E}_h\left[\frac{\partial x_j^h}{\partial q_i}\right]=\bar{x}_i+\sum_{j=1}^{n} t_j\mathbb{E}_h\left[\frac{\partial x_j^h}{\partial q_i}\right]. \tag{DM''}$$

Define $\tilde{t}_j := t_j + \mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial \bar{x}_j}\right]$. Then formula (DM') can be written

$$\sum_{j=1}^{n} \tilde{t}_j \mathbb{E}_h\left[\frac{\partial \hat{x}_j^h}{\partial q_i}\right] = \mathbb{E}_h\left[x_i^h \frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial I}\right] - \bar{x}_i.$$

Structurally, this equation is *exactly* the standard Diamond-Mirrlees formula for economies without externalities. This result was first derived by Sandmo (1975), and it implies that the optimal taxes take additive form $t_j = \tilde{t}_j + t_j^P$, where the "Pigouvian correction" is $t_j^P := -\mathbb{E}_h\left[\frac{\lambda^h}{\kappa}\frac{\partial V^h}{\partial \bar{x}_j}\right]$.

Two observations: First, the additive property demonstrates that we should *not* tax complements or subsidize substitutes for a good with a negative externality. We saw this when we considered Pigouvian taxation with type-specific lump-sum taxes, and it is notable that it continues to hold with fewer tax instruments. Second, the Pigouvian correction depends directly on the Pareto weights λ^h. Corrective taxes are now directly sensitive to the government's redistributive objective, contrary to what we found with type-specific lump-sum taxes. Intuitively, the government can no longer handle redistribution "behind the scenes" with lump-sum taxes – it compensates by placing larger corrective taxes on goods whose negative consumption externalities are borne by types with higher Pareto weights.

References

Diamond, P. A., & Mirrlees, J. A. (1971). Optimal taxation and public production ii: Tax rules. *The American Economic Review*, *61*(3), 261–278.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *The Swedish Journal of Economics*, 86–98.

Recitation 4: Chamley-Judd Revisited

Todd Lensman

February 25, 2022

Recitation Plan: Review the zero capital taxation result of Chamley (1986) and the criticism of Straub and Werning (2020)

1 Model

Consumption. The economy exists in discrete time $t \in \{0, 1, \ldots\}$ and consists of a representative agent. The agent has intratemporal preferences over consumption c_t and labor supply n_t represented by a utility function $U(c_t, n_t)$. We assume that U is such that consumption and leisure are both normal goods:

$$\frac{U_{cc}}{U_c} - \frac{U_{nc}}{U_n}, \frac{U_{cn}}{U_c} - \frac{U_{nn}}{U_n} \leq 0.$$

The agent's intertemporal utility satisfies the stationary recursion

$$V_t = W(U_t, V_{t+1}) \quad \text{and} \quad V_t = \mathcal{V}\left((U_s)_{s=t}^{\infty}\right).$$

Finally, it is helpful to define the "discount factor" applied to $t+1$ continuation utility V_{t+1} in period t when the agent is in a "steady state" with constant continuation utilities:

$$\bar{\beta}(V) := W_V\left(\bar{U}(V), V\right), \quad \text{where} \quad V = W\left(\bar{U}(V), V\right).$$

At each date t, the agent can consume c_t, supply labor n_t, save capital k_{t+1}, and purchase government bonds b_{t+1}, taking the post-tax wage w_t and the post-tax return on capital and bonds R_{t+1} as given.[1] The agent then solves

$$\max_{(c_t, n_t, k_{t+1}, b_{t+1})_{t=0}^{\infty}} \mathcal{V}\left((U_s)_{s=0}^{\infty}\right) \quad \text{subject to} \quad c_0 + k_1 + b_1 \leq w_0 n_0 + R_0 k_0 + R_0^b b_0,$$
$$c_t + k_{t+1} + b_{t+1} \leq w_t n_t + R_t(k_t + b_t) \quad t \geq 1,$$

[1] By a standard arbitrage argument, the post-tax returns on capital and bonds must be equal in every period after the initial period.

$$\lim_{t\to\infty}(k_{t+1}+b_{t+1})\prod_{s=1}^{t}R_t^{-1}=0.$$

Note that the $t=0$ budget constraint must be written separately: Initial asset holdings k_0 and b_0 are fixed, so we do not have an arbitrage argument that requires the post-tax return on capital R_0 equal the post-tax return on bonds R_0^b.

Production. The production technology is described by the production function $F(k_t, n_t)$ for the final consumption good, assumed differentiable and homogeneous of degree one. Production is competitive, so the pre-tax wage w_t^* and return on capital R_t^* are determined by marginal products:

$$w_t^* = F_n(k_t, n_t) \quad \text{and} \quad R_t^* = F_k(k_t, n_t).$$

The pre- and post-tax wage and returns are related by the identities

$$w_t = \left(1-\tau_t^n\right)w_t^* \quad \text{and} \quad R_t = (1-\tau_t)\left(R_t^*-1\right)+1. \tag{1}$$

Here τ_t^n is the tax on labor income, while τ_t is the tax on (net) asset returns.

Government. The government uses taxes $\left(\tau_t^n, \tau_t\right)_{t=0}^{\infty}$ and R_0^b to finance a stream of public spending $(g_t)_{t=0}^{\infty}$. The government's implied budget constraints are

$$\begin{aligned} g_0 + R_0^b b_0 &\le \tau_0^n w_0^* + \tau_0\left(R_0^*-1\right)k_0 + b_1, \\ g_t + R_t b_t &\le \tau_t^n w_t^* + \tau_t\left(R_t^*-1\right)k_t + b_{t+1} \quad t \ge 1. \end{aligned}$$

Following Chamley (1986), the government is also constrained by a lower bound on the after-tax rate of return: $R_t \ge (1-\bar{\tau})\left(R_t^*-1\right)+1$ for $t \ge 1$. The implied bound on the tax on net asset returns is $\tau_t \le \bar{\tau}$.

Equilibrium. An equilibrium is a tuple $\left(\left(c_t, n_t, k_{t+1}, b_{t+1}, w_t^*, R_{t+1}^*, \tau_t^n, \tau_t\right)_{t=0}^{\infty}, R_0^*, R_0^b\right)$ such that

(i) the government's budget constraint is satisfied;

(ii) pre-tax prices are determined by marginal products;

(iii) the agent chooses $(c_t, n_t, k_{t+1}, b_{t+1})_{t=0}^{\infty}$ given post-tax prices; and

(iv) the resource constraint is satisfied in each period, $c_t + g_t + k_{t+1} \le F(k_t, n_t)$.

2 Capital Taxation in the Long Run

To characterize the optimal path of taxes $\left(\tau_t^n, \tau_t\right)_{t=0}^{\infty}$, we follow the primal approach: By a standard argument, an allocation $(c_t, n_t, k_{t+1})_{t=0}^{\infty}$ can be implemented in a competitive equilibrium with taxes if and only if

$$\begin{aligned} c_t + g_t + k_{t+1} &\leq F(k_t, n_t) + (1-\delta) k_t && t \geq 0, \\ R_0 &\geq (1-\bar{\tau})(F_k(k_0, n_0) - 1) + 1, \\ \frac{\mathcal{V}_{ct}}{\mathcal{V}_{c(t+1)}} &\geq (1-\bar{\tau})(F_k(k_{t+1}, n_{t+1}) - 1) + 1 && t \geq 0, \\ \sum_{t=0}^{\infty} (\mathcal{V}_{ct} c_t + \mathcal{V}_{nt} n_t) &= \mathcal{V}_{c0}\left(R_0 k_0 + R_0^b b_0\right). \end{aligned}$$

The first set of constraints ensures feasibility, the second set of constraints ensures that the capital tax bound $R_t \geq 1$ is satisfied, and the final constraint ensures implementability. The government's problem is then to choose the allocation $(c_t, n_t, k_{t+1})_{t=0}^{\infty}$ to maximize $t=0$ utility V_0 subject to the feasibility constraints, capital tax constraints, and the implementability constraint.

We begin by recalling Chamley's (1986) result:

Theorem (Chamley, 1986, Theorem 1)**.** Let $\tilde{\Lambda}_t$ denote the Lagrange multiplier on the period t resource constraint, and let $\Lambda_t := \tilde{\Lambda}_t / \mathcal{V}_{ct}$. Suppose $c_t, k_{t+1} > 0$ for $t \geq 0$, and suppose that for $t > T$ the capital tax constraints are non-binding. Then if $\Lambda_t \to \Lambda > 0$, $R_t / R_t^* \to 1$.

Proof. With $t > T$, the first-order condition for $t+1$ capital k_{t+1} is

$$\tilde{\Lambda}_t = \tilde{\Lambda}_{t+1} R_{t+1}^* \iff \mathcal{V}_{ct} \Lambda_t = \mathcal{V}_{c(t+1)} \Lambda_{t+1} R_{t+1}^*.$$

The post-tax return R_{t+1} is defined by the agent's Euler equation:

$$\mathcal{V}_{ct} = \mathcal{V}_{c(t+1)} R_{t+1}.$$

Dividing these equations yields $R_{t+1}^* / R_{t+1} = \Lambda_t / \Lambda_{t+1} \to 1$. ■

The theorem states that if the capital tax constraint is asymptotically non-binding and the government's (normalized) marginal value of resources in period t converges to a positive constant, then the optimal capital tax τ_t must converge to zero.[2] Straub and Werning (2020)

[2]Note that if $\bar{\tau} = 1$, then the condition that the capital tax constraint is asymptotically non-binding can be

offer two key criticisms of this result: First, in the standard case in which intratemporal utility U is isoelastic and additively separable and intertemporal aggregation W is additively separable, the multiplier Λ_t need not converge to a positive value and/or the capital tax constraint may not be asymptotically non-binding. As a result, the optimal capital tax may satisfy $\tau_t = \bar{\tau}$ in all periods. The proof of this result is lengthy, and I refer you to the appendix of Straub and Werning (2020). Second, Straub and Werning (2020) show that even when Chamley's (1986) result applies, as long as intertemporal aggregation W is not additively separable, the zero long-run capital tax is also accompanied by zero long-run wealth or zero long-run labor taxation:

Theorem (Straub and Werning, 2020, Proposition 6)**.** Suppose the optimal allocation converges to an interior steady state, and suppose that for $t > T$ the capital tax constraints are non-binding. Then $\tau_t \to 0$, and if $\bar{\beta}'(V) \neq 0$ at the steady-state continuation utility V, then either

(i) private wealth converges to zero, $a_t := k_t + b_r \to 0$; or

(ii) the allocation converges to the first-best, with $\tau_t^n \to 0$.

Proof. The proof follows from an examination of the first-order conditions to the government's problem. First, we define notation from the agent's preferences: Given the optimal allocation $(c_t, n_t, k_{t+1})_{t=0}^{\infty}$, define the period-$t$ discount rate $\beta_t := \prod_{s=0}^{t-1} W_V(U_s, V_{s+1})$. Then using the intertemporal recursion $\mathcal{V}_{ct} = \beta_t W_{Ut} U_{ct}$ and $\mathcal{V}_{nt} = \beta_t W_{Ut} U_{nt}$, the implementability constraint can be written in the more familiar form

$$\sum_{t=0}^{\infty} \beta_t W_{Ut} (U_{ct} c_t + U_{nt} n_t) = W_{U0} U_{c0} \left(R_0 k_0 + R_0^b b_0\right).$$

The government's problem can then be stated

$$\max_{(V_t, c_t, n_t, k_{t+1})_{t=0}^{\infty}, R_0, R_0^b} V_0 \tag{2}$$

$$\text{subject to} \tag{3}$$

$$V_t = W(U(c_t, n_t), V_{t+1}) \qquad t \geq 0, \tag{4}$$

$$c_t + g_t + k_{t+1} \leq F(k_t, n_t) \qquad t \geq 0, \tag{5}$$

$$\sum_{t=0}^{\infty} \beta_t W_{Ut} (U_{ct} c_t + U_{nt} n_t) = W_{U0} U_{c0} \left(R_0 k_0 + R_0^b b_0\right), \tag{6}$$

removed.

$$\tau_t \le \bar{\tau} \qquad\qquad t \ge 0. \tag{7}$$

Let $\beta_t \nu_t$ be the multiplier on the period-t recursion constraint, let $\beta_t \lambda_t$ be the multiplier on the period-t resource constraint, and let μ be the multiplier on the implementability constraint. Then for $t > T$, the first-order conditions for V_{t+1}, c_t, n_t, and k_{t+1} are

$$\begin{aligned}
(V_{t+1}) \quad & 0 = -\nu_t + \nu_{t+1} - \mu A_{t+1}, \\
(c_t) \quad & 0 = \nu_t W_{Ut} U_{ct} - \mu W_{Ut}\left(U_{ct} + U_{cct}c_t + U_{nct}n_t\right) - \mu B_t U_{ct} - \lambda_t, \\
(n_t) \quad & 0 = \nu_t W_{Ut} U_{nt} - \mu W_{Ut}\left(U_{nt} + U_{cnt}c_t + U_{nnt}n_t\right) - \mu B_t U_{nt} + \lambda_t F_{nt}, \\
(k_{t+1}) \quad & 0 = -\lambda_t + \lambda_{t+1} W_{Vt} F_{k(t+1)},
\end{aligned}$$

where we have used the assumption that the capital tax constraints are non-binding and

$$\begin{aligned}
A_{t+1} &:= \frac{1}{\beta_{t+1}} \frac{\partial}{\partial V_{t+1}} \sum_{s=0}^{\infty} \beta_s W_{Us}\left(U_{cs}c_s + U_{ns}n_s\right) \\
B_t &:= \frac{1}{\beta_t} \sum_{s=0}^{\infty} \frac{\partial\left(\beta_s W_{U_s}\right)}{\partial U_t}\left(U_{cs}c_s + U_{ns}n_s\right).
\end{aligned}$$

Now we suppose that $(c_t, n_t, k_{t+1}) \to (c, n, k)$, which implies that utilities, continuation values, and their derivatives also converge. Assets $a_t := k_t + b_t$ converge, and the limit can be found by using a period $t+1$ version of the implementability constraint:

$$a_{t+1} = \frac{\sum_{s=t+1}^{\infty} \beta_s W_{U_s}\left(U_{cs}c_s + U_{ns}n_s\right)}{W_{U(t+1)}U_{c(t+1)}\beta_{t+1}R_{t+1}} \to \frac{U_c c + U_n n}{\left(1 - \bar{\beta}(V)\right) U_c R} =: a,$$

where the limit holds because the sum is asympotically a geometric series. A similar argument implies $A_{t+1} \to A$ and $B_t \to B$, where

$$A = \frac{\bar{\beta}'(V)}{\bar{\beta}(V)} W_U U_c R a.$$

Taking the limits for all allocation variables in the first-order conditions, we have

$$\begin{aligned}
(V) \quad & 0 = -\nu_t + \nu_{t+1} - \mu A, \\
(c) \quad & 0 = \nu_t - \mu\left(1 + \frac{U_{cc}c}{U_c} + \frac{U_{nc}n}{U_c}\right) - \mu \frac{B}{W_U} U_c - \frac{\lambda_t}{W_U U_c}, \\
(n) \quad & 0 = \nu_t - \mu\left(1 + \frac{U_{cn}c}{U_n} + \frac{U_{nn}n}{U_n}\right) - \mu \frac{B}{W_U} + \lambda_t \frac{F_n}{W_U U_n},
\end{aligned}$$

$$(k) \quad 0 = -\lambda_t + \lambda_{t+1} \bar{\beta}(V) F_k.$$

Making use of the conditions for V, c, and k, we find

$$\bar{\beta}(V) F_k - 1 = \frac{\lambda_t}{\lambda_{t+1}} - 1 = -\frac{W_U U_c}{\lambda_{t+1}} \mu A.$$

To prove the theorem, we first show that capital taxes are indeed zero, $\bar{\beta}\left(\bar{V}\right) F_k = 1$. If $A = 0$ or $\mu = 0$, then this is immediate from the equation above. Otherwise, the V condition requires that $\nu_t \to \pm\infty$, and hence that $\lambda_t \to \pm\infty$. We again recover $\bar{\beta}\left(\bar{V}\right) F_k = 1$ using the equation above.

We can complete the argument by showing that $a \neq 0$ and $\bar{\beta}'(V) \neq 0$ imply $\tau_t^n = 0$. Using the conditions for c and n, the labor tax satisfies

$$\lambda_t \tau^n = \mu \frac{W_U U_n}{F_n} \left[\frac{U_{cc} c}{U_c} + \frac{U_{nc} n}{U_c} - \left(\frac{U_{cn} c}{U_n} + \frac{U_{nn} n}{U_n} \right) \right].$$

To see that $\tau^n = 0$, note that $\mu = 0$ implies that the economy is first-best, which immediately implies $\tau_t^n = 0 \ \forall t \geq 0$. Suppose instead that $\mu \neq 0$. If $\lambda_t \to \pm\infty$, then the equation above immediately implies $\tau^n = 0$. Suppose instead that $\lambda_t \to \lambda \in \mathbb{R}$.[3] This implies $\nu_t \to \nu$, and hence that $A = 0$. But this is a contradiction since we assume that $a, \bar{\beta}'(V) \neq 0$. ■

This result suggests caution in interpreting Chamley's (1986) zero capital taxation result away from the "knife-edge" case of additively separable intertemporal aggregation: Though zero capital taxation may be optimal in the long run, this long run must also feature either zero labor taxation (symmetric treatment of capital and labor) or zero private wealth!

References

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, 607–622.

Straub, L., & Werning, I. (2020). Positive long-run capital taxation: Chamley-judd revisited. *American Economic Review*, *110*(1), 86–119.

[3] This is the only other possibility, because zero capital taxation implies $\frac{\lambda_t}{\lambda_{t+1}} \to 1$.

Recitation 5: Nonlinear Taxation I

Todd Lensman

March 4, 2022

Recitation Plan: Mathematically formulate the nonlinear income taxation problem with two types, and discuss the generalization of Stiglitz (1982) with endogenous wages

1 Model

Consumption. The economy has a measure μ_i of agents of type $i \in \{1,2\}$, where $\mu_1 + \mu_2 = 1$. Each agent of type i has preferences over consumption c_i and labor n_i given by the utility function $u^i(c_i, n_i)$, assumed twice continuously differentiable, strictly concave, strictly increasing in c_i, and strictly decreasing in n_i. Given an income tax schedule T and a pre-tax wage w_i, a type i agent solves

$$\max_{c_i, n_i} u^i(c_i, n_i) \quad \text{subject to} \quad c_i \leq w_i n_i - T(w_i n_i).$$

Production. The production technology is described by the production function $F(n_1, n_2)$ for the final consumption good, assumed twice continuously differentiable and homogeneous of degree one. Production is competitive, so the pre-tax wages (w_1, w_2) are determined by the marginal products:

$$w_i = F_{n_i}(\mu_1 n_1, \mu_2 n_2),$$

or equivalently

$$w_2 = f'(n) \quad \text{and} \quad w_1 = f(n) - n f'(n),$$

where $n := \mu_2 n_2 / \mu_1 n_1$ and $f(n) := F(1, n)$.

Government. The government uses the income tax schedule T to finance exogenous government expenditures g and to redistribute. The government's budget constraint is

$$g \leq \mu_1 T(w_1 n_1) + \mu_2 T(w_2 n_2).$$

Equilibrium. An equilibrium is a tuple $\left((c_i, n_i)_{i=1,2}, T\right)$ such that

(i) the government's budget constraint is satisfied;

(ii) pre-tax wages are determined by marginal products;

(iii) each type i agent chooses (c_i, n_i) given the pre-tax wage w_i and the tax schedule T; and

(iv) the resource constraint is satisfied in each period, $\mu_1 c_1 + \mu_2 c_2 + g \leq F(\mu_1 n_1, \mu_2 n_2)$.

2 Special Case: Linear Production

First consider the case in which the production function F is linear: $F(\mu_1 n_1, \mu_2 n_2) = w_1 \mu_1 n_1 + w_2 \mu_2 n_2$, where $w_2 > w_1$. Additionally assume the following single-crossing condition:

$$\mathrm{MRS}^2(c,y) := -\frac{1}{w_2}\frac{u_n^2(c, y/w_2)}{u_c^2(c, y/w_2)} < -\frac{1}{w_1}\frac{u_n^1(c, y/w_1)}{u_c^1(c, y/w_1)} =: \mathrm{MRS}^1(c,y) \quad \forall (c,y) \gg 0.$$

This condition implies that the indifference curves for a type 2 agent are flatter than those for a type 1 agent in (y,c)-space, and it will imply that type 2 agents will earn higher incomes under a Pareto efficient income tax.

The Pareto efficient income taxation problem is as follows: Choose the tax schedule T to maximize the equilibrium utility of type 2 agents, subject to the government's budget constraint and the constraint that type 1 agents achieve utility $\bar{u}^1$ in equilibrium. This is a difficult problem! As stated, the "choice variable" is an infinite-dimensional object (the tax schedule T). To simplify, we make use of a change of variables known in mechanism design as the Revelation Principle: Fix a tax schedule T, and let (c_i, n_i) denote the consumption bundle chosen by an agent of type i in equilibrium. Since agents optimize, these consumption bundles must satisfy the "incentive constraints"

$$u^i(c_i, n_i) \geq u^i\left(c_{i'}, \frac{w_{i'}}{w_i} n_{i'}\right) \quad i \neq i'.$$

In words, a type i agent must weakly prefer her own bundle to that of a type i' agent; otherwise, she would have chosen the type i' agent's bundle. Conversely, suppose that we find consumption bundles $(c_i, n_i)_{i=1,2}$ that satisfy the incentive constraints above as well as the resource constraint. By the single-crossing condition, we must have $w_2 n_2 > w_1 n_1$, and we can define an income tax schedule by

$$T(y) := \begin{cases} w_i n_i - c_i & \text{if } y = w_i n_i \text{ for } i = 1, 2, \\ \infty & \text{else.} \end{cases}$$

It is straightforward to verify that if the consumption bundles $(c_i, n_i)_{i=1,2}$ satisfy the incentive compatibility conditions, then an agent of type i will choose bundle (c_i, n_i) when confronted with the income tax schedule T. As a result, instead of formulating the Pareto efficiency problem as an optimization problem over tax schedules, we can instead choose an allocation $(c_i, n_i)_{i=1,2}$ subject to incentive compatibility constraints:

$$\begin{aligned} \max_{(c_i, n_i)_{i=1,2}} u^2(c_2, n_2) \quad \text{subject to} \quad & u^1(c_1, n_1) \geq \bar{u}^1, \\ & u^i(c_i, n_i) \geq u^i\left(c_{i'}, \frac{w_{i'}}{w_i} n_{i'}\right) \qquad i \in \{1, 2\}, i \neq i', \\ & \mu_1 c_1 + \mu_2 c_2 + g \leq F(\mu_1 n_1, \mu_2 n_2). \end{aligned}$$

We can characterize properties of the solution using first-order conditions. Let $\eta > 0$ denote the multiplier on the utility constraint, let $\lambda_i \geq 0$ denote the multiplier on type i's incentive constraint, and let $\gamma > 0$ denote the multiplier on the resource constraint. The first-order conditions are

$$\begin{aligned} (c_1) \quad & 0 = (\eta + \lambda_1) u_c^1(c_1, n_1) - \lambda_2 u_c^2\left(c_1, \frac{w_1}{w_2} n_1\right) - \gamma \mu_1, \\ (n_1) \quad & 0 = (\eta + \lambda_1) u_n^1(c_1, n_1) - \frac{w_1}{w_2} \lambda_2 u_n^2\left(c_1, \frac{w_1}{w_2} n_1\right) + \gamma w_1 \mu_1, \\ (c_2) \quad & 0 = (1 + \lambda_2) u_c^2(c_2, n_2) - \lambda_1 u_c^1\left(c_2, \frac{w_2}{w_1} n_2\right) - \gamma \mu_2, \\ (n_2) \quad & 0 = (1 + \lambda_2) u_n^2(c_2, n_2) - \frac{w_2}{w_1} \lambda_1 u_n^1\left(c_2, \frac{w_2}{w_1} n_2\right) + \gamma w_2 \mu_2, \end{aligned}$$

Using these conditions, we can show by contradiction that we cannot have $\lambda_1, \lambda_2 > 0$, i.e., at most one incentive constraint can bind. We will suppose that $\bar{u}^1$ is sufficiently high so that

$\lambda_2 > 0 = \lambda_1$: The government wishes to redistribute from high-wage type 2 to low-wage type 1 agents, so the type 2 agents' incentive constraint would not be satisfied in the first-best allocation. In this case, the type 2 first-order conditions are

$$\begin{aligned}
(c_2) \quad & 0 = (1+\lambda_2)u_c^2(c_2, n_2) - \gamma\mu_2, \\
(n_2) \quad & 0 = (1+\lambda_2)u_n^2(c_2, n_2) + \gamma w_2 \mu_2.
\end{aligned}$$

Dividing, we find

$$\mathrm{MRS}^2(c_2, w_2 n_2) = 1.$$

The marginal rate of substitution between consumption and labor is equalized with the wage, so we recover the "no distortion at the top" result: Type 2 agents must face a marginal tax rate of zero at their equilibrium income $w_2 n_2$. Similarly, type 1's first-order conditions are

$$\begin{aligned}
(c_1) \quad & 0 = \eta u_c^1(c_1, n_1) - \lambda_2 u_c^2\left(c_1, \frac{w_1}{w_2} n_1\right) - \gamma\mu_1, \\
(n_1) \quad & 0 = \eta u_n^1(c_1, n_1) - \frac{w_1}{w_2}\lambda_2 u_n^2\left(c_1, \frac{w_1}{w_2} n_1\right) + \gamma w_1 \mu_1.
\end{aligned}$$

Dividing yields

$$\begin{aligned}
\mathrm{MRS}^1(c_1, w_1 n_1) &= \frac{1 - \left(\frac{\lambda_2}{w_2} u_n^2\left(c_1, \frac{w_1}{w_2} n_1\right)\right)/(\gamma\mu_1)}{1 + \left(\lambda_2 u_c^2\left(c_1, \frac{w_1}{w_2} n_1\right)\right)/(\gamma\mu_1)} \\
&= \mathrm{MRS}^2(c_1, w_1 n_1) + \frac{1 - \mathrm{MRS}^2(c_1, w_1 n_1)}{1+\nu},
\end{aligned}$$

where

$$\nu := \frac{\lambda_2 u_c^2\left(c_1, \frac{w_1}{w_2} n_1\right)}{\gamma\mu_1}.$$

Rearranging the equation above and making use of the single-crossing assumption, we have

$$\begin{aligned}
(1+\nu)\mathrm{MRS}^1(c_1, w_1 n_1) &= \nu \mathrm{MRS}^2(c_1, w_1 n_1) + 1 \\
&< \nu \mathrm{MRS}^1(c_1, w_1 n_1) + 1 \\
\Rightarrow \mathrm{MRS}^1(c_1, w_1 n_1) &< 1.
\end{aligned}$$

Thus the marginal rate of substitution between consumption and labor is below the wage w_1, so type 1 agents face a positive marginal tax rate at their equilibrium income $w_1 n_1$.

3 Nonlinear Production

We now relax the assumption that the production function F is linear. In this case, when solving the Pareto efficiency problem we must be careful to incorporate the general equilibrium determination of the wages (w_1, w_2). The equilibrium relative wage is determined by

$$\frac{w_1}{w_2} = \frac{f(n) - nf'(n)}{f'(n)} =: \phi\left(\frac{n_2}{n_1}\right).$$

The incentive constraints can then be written

$$\begin{aligned} u^1(c_1, n_1) &\geq u^1\left(c_2, \frac{n_2}{\phi(n_2/n_1)}\right), \\ u^2(c_2, n_2) &\geq u^2\left(c_1, \phi\left(\frac{n_2}{n_1}\right) n_1\right). \end{aligned}$$

Assuming $w_2 > w_1$ and $\lambda_2 > 0 = \lambda_1$ at the solution to the modified Pareto efficiency problem, we find the first-order conditions

$$\begin{aligned} &(c_1) \quad 0 = \eta u_c^1(c_1, n_1) - \lambda_2 u_c^2(c_1, n_1\phi) - \gamma\mu_1, \\ &(n_1) \quad 0 = \eta u_n^1(c_1, n_1) - \lambda_2 u_n^2(c_1, n_1\phi)\left(\phi - \frac{n_2}{n_1}\phi'\right) + \gamma F_1 \mu_1, \\ &(c_2) \quad 0 = (1 + \lambda_2) u_c^2(c_2, n_2) - \gamma\mu_2, \\ &(n_2) \quad 0 = (1 + \lambda_2) u_n^2(c_2, n_2) - \lambda_2 u_n^2(c_1, n_1\phi)\phi' + \gamma F_2 \mu_2. \end{aligned}$$

Dividing the first-order conditions for type 2, we find

$$\mathrm{MRS}^2(c_2, w_2 n_2) = 1 - \frac{\lambda_2 u_n^2(c_1, n_1\phi)\phi'}{\gamma w_2 \mu_2}.$$

Provided that $\phi' > 0$, we then find that the marginal rate of substitution between consumption and labor is strictly greater than the wage for type 2 agents, implying a *negative* marginal tax rate at the equilibrium income $w_2 n_2$. We can similarly use the first-order conditions for type 1 agents to show that they continue to face a positive marginal tax rate at the optimum, and that this marginal tax rate is decreasing in the elasticity of substitution between both types of labor in production.

The intuition for these results is as follows: When type 1 labor and type 2 labor are not perfectly substitutable in production, the relative quantities employed will affect the marginal product of each factor. By subsidizing employment of type 2 agents, the government raises the relative wage of type 1 agents, making it less attractive for a type 2 agent to mimick a type 1 agent. The government can then allow type 1 agents to work and consume more so as to increase their utilities while ensuring that the type 2 incentive constraint is relaxed relative to the fixed wage benchmark. In this sense, "predistribution" through manipulation of factor prices can benefit the government, an idea considered again in Naito (1999).

References

Naito, H. (1999). Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics*, *71*(2), 165–188.

Stiglitz, J. E. (1982). Self-selection and pareto efficient taxation. *Journal of public economics*, *17*(2), 213–240.

Recitation 6: Nonlinear Taxation II

Todd Lensman

March 11, 2022

Recitation Plan: Review the heuristic derivation of the optimal nonlinear income tax formula in Saez (2001)

1 Model

Consumption. The economy has a unit measure of heterogeneous agents, where the density of agents of type $\theta \in \mathbb{R}_+$ is $f(\theta)$. An agent of type θ has preferences over consumption c and income (effective labor) y given by the utility function $u(c, y, \theta)$, assumed twice continuously differentiable, strictly concave, strictly increasing in c and θ, and strictly decreasing in y. We also assume that u satisfies the single-crossing condition, so that the marginal rate of substitution

$$-\frac{u_y(c, y, \theta)}{u_c(c, y, \theta)}$$

is decreasing in θ. Given an income tax schedule T, an additional linear tax τ, and an additional lump-sum grant I, a type θ agent solves

$$\max_{c,y} u(c, y, \theta) \quad \text{subject to} \quad c \leq y - T(y) - (1-\tau) y + I.$$

Let $y(\theta; T, \tau, I)$ denote the resulting income choice, which is weakly increasing in θ and strictly increasing wherever $y \mapsto T(y) - (1-\tau) y$ is smooth. Suppose that T is smooth, and define the uncompensated elasticity, the income effect, and the compensated elasticity as follows:

$$\begin{aligned}
\varepsilon^u(y; T) &:= \frac{1 - T'(y)}{y} \frac{\partial y(\theta; T, 0, 0)}{\partial (1-\tau)}, \\
\eta(y; T) &:= -\left(1 - T'(y)\right) \frac{\partial y(\theta; T, 0, 0)}{\partial I}, \\
\varepsilon^c(y; T) &:= \varepsilon^u(y; T) + \eta(y; T),
\end{aligned}$$

where $y = y(\theta; T, 0, 0)$. In the remainder of the note, I set $\tau = I = 0$ and write $y = y(\theta; T)$.

Production. The production technology is such that income is transformed one-for-one into consumption.

Government. The government uses the income tax schedule T to finance exogenous government expenditures E and to redistribute. The government's budget constraint is

$$E \leq \int_0^\infty T(y) H(dy; T),$$

where the endogenous income distribution H satisfies $H(y; T) := F\left(y^{-1}(\theta; T)\right)$. The government's redistributive objective is described by the welfare function

$$\int_0^\infty G(u(c(\theta), y(\theta))) f(\theta) d\theta.$$

Equilibrium. An equilibrium is a tuple $\left((c(\theta), y(\theta))_{\theta \in \mathbb{R}_+}, T\right)$ such that

(i) the government's budget constraint is satisfied;

(ii) each type θ agent chooses $(c(\theta), y(\theta))$ given the tax schedule T.

2 Heuristic Derivation of the Diamond-Saez Formula

The government's task is to choose a tax schedule T to maximize its welfare function, subject to its budget constraint. Let φ denote the multiplier on the resource constraint, which can be interpreted as the value of public funds at the optimum. To derive an optimality condition that characterizes the optimal tax schedule T, we use a variational argument: We consider small changes to the optimal tax schedule, and we argue (just as in finite-dimensional calculus) that as these changes become small, they must have no first-order effect on the government's Lagrangian (penalized objective function). In particular, for each income level y^*, we consider a variation that increases the marginal tax rate by $d\tau$ for all incomes between y^* and $y^* + dy^*$. The effect of this variation is to increase the marginal tax rate for all income levels $[y^*, y^* + dy^*]$ while increasing only the average tax rate for all income levels $[y^* + dy^*, \infty)$. The effect of this variation can be divided into three components.

Mechanical Effect. Holding behavioral responses constant, the variation has direct effects on (i) the government's budget constraint and (ii) the government's objective. In particular, every agent with income above y^*+dy^* pays additional income taxes $d\tau dy^*$. Any agent with income $y \in (y^* + dy^*)$ pays additional income taxes $d\tau (y - y^*)$. The total effect on the government's revenues is

$$d\tau \int_{y^*}^{y^*+dy^*} (y-y^*)h(y)\,dy + d\tau dy^* \int_{y^*+dy^*}^{\infty} h(y)\,dy,$$

where $h(y) := H'(y;T)$ is the income density at the optimal tax, which is assumed to exist. Note that the first term is of smaller order than the second term in $d\tau dy^*$: It tends to zero as $d\tau, dy^* \to 0$ when renormalized by $d\tau dy^*$, while the second term tends to a non-zero limit when renormalized by $d\tau dy^*$. We are only concerned about first-order changes around the optimal tax, so we can freely ignore the first term in our subsequent analysis. Similarly, the effect on the government's objective is

$$-d\tau dy^* \int_{y^*+dy^*}^{\infty} g(y)h(y)\,dy,$$

where I have ignored the second-order term and defined

$$g(y) := \frac{G'(u(c(\theta), y(\theta)))u_c(c(\theta), y(\theta))}{\varphi}$$

for $y = y(\theta;T)$. Combining the previous two terms, the first-order mechanical effect is

$$M := d\tau dy^* \int_{y^*}^{\infty} (1 - g(y))h(y)\,dy.$$

Substitution Effect. We now consider the first of two behavioral responses: substitution effects for agents who initially selected incomes $y \in [y^* + dy^*]$. Any such agent now faces a higher marginal tax rate $d\tau$ and adjusts her income according to the compensated elasticity ε^c:

$$dy = -d\tau \frac{y}{1-T'(y)} \varepsilon^c(y).$$

By the Envelope Theorem, this behavioral adjustment has no first-order effect on the agent's utility, so it does not impact the government's objective. However, it does have a first-order effect ("fiscal externality") on the government's budget constraint:

$$\int_{y^*}^{y^*+dy^*} T'(y)\left[-d\tau \frac{y}{1-T'(y)}\varepsilon^c(y)\right]h(y)\,dy.$$

For dy^* small, the total effect of substitution on the budget constraint is

$$S := -d\tau dy^* \frac{y^* T'(y^*)}{1-T'(y^*)}\varepsilon^c(y^*)h(y^*).$$

Income Effect. The second behavioral response results from income effects for agents who initially selected incomes $y \in [y^* + dy^*, \infty)$. Any such agent faces the same marginal tax rate $T'(y)$, but with her after-tax income reduced by $d\tau dy^*$. She then adjusts her income according to the income effect η:

$$dy = -d\tau dy^* \frac{\eta(y)}{1-T'(y)}.$$

Again, this adjust has first-order effect only on the government's budget constraint:

$$I := -d\tau dy^* \int_{y^*}^{\infty} \frac{T'(y)}{1-T'(y)}\eta(y)h(y)\,dy.$$

Optimality. At an optimal tax schedule T, the mechanical, substitution, and income effects must sum to zero:

$$0 = M + S + I.$$

Replacing $y = y^*$, this equation can be written

$$\frac{T'(y)}{1-T'(y)}\varepsilon^c(y)yh(y) = \int_y^{\infty}(1-g(\tilde{y}))h(\tilde{y})\,d\tilde{y} - \int_y^{\infty}\frac{T'(\tilde{y})}{1-T'(\tilde{y})}\eta(\tilde{y})h(\tilde{y})\,d\tilde{y}. \tag{1}$$

Remarkably, we can differentiate this equation to derive the Pareto optimality inequality of Werning (2007):

Proposition 2.1 (Werning, 2007)**.** A tax schedule T is optimal for some income-dependent

Pareto weights $g(z) \geq 0$ only if

$$-\frac{T'(y)}{1-T'(y)}\varepsilon^c(y)\left[\frac{\partial \log(yh(y))}{\partial \log(y)} + \frac{\partial \log\left(\frac{T'(y)}{1-T'(y)}\varepsilon^c(y)\right)}{\partial \log(y)} + \frac{\eta(y)}{\varepsilon^c(y)}\right] \leq 1.$$

Alternatively, we can view (1) as an ordinariy differential equation with variable coefficients:

$$\underbrace{\frac{T'(y)}{1-T'(y)}\eta(y)h(y)}_{\dot{K}(y)} = \underbrace{\frac{\eta(y)}{\varepsilon^c(y)y}}_{D(y)}\left[\underbrace{\int_y^\infty (1-g(\tilde{y}))h(\tilde{y})d\tilde{y}}_{C(y)} - \underbrace{\int_y^\infty \frac{T'(\tilde{y})}{1-T'(\tilde{y})}\eta(\tilde{y})h(\tilde{y})d\tilde{y}}_{K(y)}\right].$$

Integrating with the boundary condition $\lim_{y\to\infty} K(y) = 0$ yields

$$\begin{aligned} K(y) &= -\int_y^\infty D(\tilde{y})C(\tilde{y})\exp\left(-\int_y^{\tilde{y}} D(z)dz\right)d\tilde{y} \\ &= -\int_y^\infty C'(\tilde{y})\exp\left(-\int_y^{\tilde{y}} D(z)dz\right)d\tilde{y} - C(y), \end{aligned}$$

where the second line follows by integrating by parts. A final differentiation gives the standard Diamond-Saez formula:

$$\frac{T'(y)}{1-T'(y)} = \frac{1-H(y)}{\varepsilon^c(y)yh(y)}\int_y^\infty (1-g(\tilde{y}))\exp\left(\int_y^{\tilde{y}} \frac{\eta(z)}{\varepsilon^c(z)}\frac{dz}{z}\right)\frac{h(\tilde{y})d\tilde{y}}{1-H(y)}. \tag{2}$$

References

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies*, *68*(1), 205–229.

Werning, I. (2007). *Pareto efficient income taxation*. mimeo, MIT.