Constraints on Income Distribution and Production Efficiency In Economies with Ramsey Taxation

Charles Blackorby and Sushama Murty

No 908

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Constraints on Income Distribution and Production Efficiency In Economies with Ramsey Taxation*

Charles Blackorby and Sushama Murty

May 2009

Charles Blackorby: Department of Economics, University of Warwick: c.blackorby@warwick.ac.uk

Sushama Murty: Department of Economics, University of Warwick: s.murty@warwick.ac.uk

Abstract

We study the link between second-best production efficiency and the constraints on income distribution imposed by private ownership of firms in economies with Ramsey taxation. We review the result of Dasgupta and Stiglitz [1972], Mirrlees [1972], Hahn [1973], and Sadka [1977] about firm-specific profit taxation leading to second-best production efficiency. Problems in the proofs of this result in these papers have been identified by Reinhorn [2005]. We provide an alternative, and with some hope a more intuitive, proof of this result. The mechanism employed in our proof is also used to show second-best production efficiency under some configuarations of private ownership without any (or at best, uniform) profit taxation. The results obtained raise questions about the genericity of the phenomenon of second-best production inefficiency and about recovering social shadow prices in such economies.

Journal of Economic Literature Classification Number:H21

Keywords: Ramsey taxation, production inefficiency, general equilibrium, private ownership.

Constraints on Income Distribution and Production Efficiency In Economies with Ramsey Taxation

1. Introduction.

Diamond and Mirrlees (DM) [1971] revisited the problem first posed by Ramsey [1927] about alternative policy instruments that can be employed when there are informational constraints on the implementation of the second-welfare theorem. They showed that, when personalized lump-sum transfers are not available to the government as redistributive devices, commodity taxation can be employed as an alternative, albeit second-best, means of redistribution. In an economy with commodity taxation and constant returns to scale in production they characterized the second-best optimal commodity taxes and demonstrated, under certain weak conditions, that all the second-best optimal allocations are production efficient. The latter result was striking for two reasons. Firstly, it provided a counter example to the nihilistic claims of Lipsey and Lancaster [1956] and secondly it justified the use of producer prices as perfect proxies for the generally unobservable shadow prices of commodities in such economies. As demonstrated in Little and Mirrlees [1974] and Dreze and Stern [1987], in such economies, the producer prices can be used to evaluate and choose among marginal public sector projects.

The income distribution scheme underlying the DM model is very simple. Consumers derive incomes from their endowments.[1] When this model is extended to allow for decreasing returns-to-scale in production, and hence the existence of positive profits, it is known that, as long as the government can tax away pure profits at one-hundred percent and redistribute the proceeds in the form of a demogrant, the production efficiency result of DM continues to hold (for models with this assumption, see Guesnerie [1995]).

If constraints on redistribution are imposed this production efficiency result is jeopardized. This is discussed in a series of papers starting with Dasgupta and Stiglitz [1972], Mirrlees [1972], and Hahn [1973], which examine the DM model when, as in a Arrow-Debreu economy, consumers can hold ownership shares in the pure profits of firms and hence derive profit incomes in addition to the demogrant and their endowment incomes. However, these papers demonstrate that, as long as the government can implement firm-specific profit taxation, production efficiency continues to be desirable at all second-best allocations.

The assumption that government can implement hundred percent profit taxation is standard in the general equilibrium literature on Ramsey taxation (see Guesnerie [1995]). It is however hard to believe that the governments of mixed economies have strong enough

[1] The DM results hold also if the model is extended to allow for a demogrant (also known as a poll tax/transfer or a uniform lump-sum tax/transfer) financed out of the receipts from commodity taxation.

taxation powers to implement hundred percent profit taxation or firm specific profit taxes.[2] Presumably these assumptions continue to be made in most theoretical works in this area because of the tractability that they lend to general equilibrium analyses of Ramsey taxes or because they preserve the production efficiency results of DM, which are so convenient for cost-benefit tests.

The aim of this paper is to explore the link between income-distribution schemes involving private-ownership of firms that are available in economies with Ramsey taxes and production efficiency at the second-best allocations of such economies. First, we note that the precise mechanism ensuring production efficiency when profit taxation is available is not clear in the earlier works and that the errors in the proofs of Hahn and Mirrlees pointed out by Reinhorn [2005] and Sadka [1977] seem robust. Hence we provide an alternative proof that is simpler, more intuitive, and makes explicit the precise mechanism underlying the results of Dasgupta and Stiglitz, Mirrlees, and Hahn. At the same time, it also reveals why this particular mechanism fails to generate the DM production efficiency result in most private ownership economies where the government does not have recourse to one-hundred percent profit taxation or firm-specific profit taxation. In addition, we identify certain structures of private ownership where production efficiency holds at all second-best equilibria when there is no (or at best uniform) profit taxation. The one-hundred percent profit taxation with a demogrant case (which, incidentally, is equivalent to a private ownership economy, where all consumers have the same shares in each firm) turns out to be a special case of such economies.

These results have important implications for cost-benefit analysis, as they imply that in more realistic cases, where there are more complicated income distribution rules than those studied by DM (for instance, income distributions involving private ownership of firms) and constraints on government's taxation powers, the production efficiency result of DM fails. Producer prices can no longer be used as proxies for shadow prices. This leads to further interesting and important questions: (i) how likely is it to encounter economies where production efficiency fails, (ii) if production efficiency does not hold globally at *all* second-best allocations of a given economy, then what is the size of the subset of second-best points where it does hold–is this subset negligible, (iii) if production efficiency fails to hold at the second-best, then what is the relation between shadow prices and the prices that can be observed in the real world, and (iv) how can we recover shadow prices from the data if we are in a world where production efficiency is not true? These, however, are questions to be addressed in future projects and this paper is a necessary step towards their resolution.

[2] Hahn however, argues that the information to do so is available.

Section 2 lays down the model and the definition of second-best production efficiency that we employ in this paper. Section 3 discusses the earlier results on production efficiency, some shortcomings in the proofs of these results, and the alternative strategy that is proposed in this paper. Section 4 states some preliminary lemmas. Section 5 takes up the issue of production efficiency with firm-specific taxation. Section 6, employing the strategy of Section 5, explores the structures of private ownership that yield second-best efficiency without one hundred percent profit taxation. Section 7, provides an example of production inefficiency, and Section 8 concludes.

2. A Working Definition of Second-Best Production Efficiency and the Model

In a second-best welfare maximization problem of the government, the shadow prices (the social value) of goods in the economy are indicated by the values of the Lagrange multipliers of the resource constraints at the optimum. If consumers' preferences satisfy local nonsatiation and there is public sector production where the government is free to choose any point in the public sector technology then, at a second-best Pareto optimum, the public sector production vector lies on the frontier of the public sector technology and the shadow prices of the resources are proportional to the shadow prices in the public sector.

Under standard regularity assumptions, for every (frontier) public-sector production vector, there is a price vector that rationalizes that choice as a profit maximizing one for the public sector. So a behaviorally unconstrained choice of a production vector by the public sector is equivalent to it responding competitively to a price vector that may be different from the private sector producer price vector.

Thus, as in all the previous papers, the working definition of second-best production efficiency that we adopt in this paper is the proportionality of producer prices in the private and public sectors at a second-best Pareto optimum. This is equivalent to the aggregate production vector lying on the frontier of the aggregate technology of the economy at the second-best, where the aggregate technology of the economy is defined as the vector sum of the technologies of the private firms and the public sector.

There are N commodities indexed by k, H consumers indexed by h, $I + 1$ firms indexed by i. The firm indexed by 0 is the government firm, while the rest of the firms are private sector firms.

For every firm i, the technology is denoted by Y^i. The aggregate technology is $Y = \sum_i Y^i$. We also define the aggregate technology of the private sector as $Y^c := \sum_{i \neq 0} Y^i$. For every firm i, the technology is denoted by Y^i. For all i, let $B^i := \{\rho \in$

$\mathbf{R}^N_+ \setminus \{0^N\}|\ \rho \cdot y$ is bounded from above for all $y \in Y^i\}$. For all $i \neq 0$ we define the profit function $\pi^i : B^i \to \mathbf{R}$ with image

$$\pi^i(\rho) = \sup_y \{\rho \cdot y | y \in Y^i\}. \tag{2.1}$$

The supply vector is the the solution mapping of (2.1) $y^i : B^i \to \mathbf{R}^N$ with image $y^i(\rho)$. Private firms face a common producer price vector denoted by $p \in \cap_{i \neq 0} B^i$. The price vector faced by the public sector is denoted by $p^0 \in B^0$.

For all h, the *net* consumption set of consumer h is $X^h \subset \mathbf{R}^N$ and the preferences over X^h are represented by continuous, quasi-concave, and locally nonsatiated utility functions $u^h : X^h \to \mathbf{R}$ with images $u^h(x^h)$.[3]

Consumers face consumer prices $q \in \mathbf{R}^N_+$. For all h, the net (of endowment) income is denoted by m^h. Let the mapping $x^h : \mathbf{R}^N_+ \times \mathbf{R} \to \mathbf{R}^N$ with image $x^h(q, m^h)$ denote the Marshallian demand vector of consumer h. Every consumer h receives a demogrant, which is be denoted by $m \in \mathbf{R}$.

If the profits of the private firms are distributed to consumers, then there is an exogeneous $H \times I$-dimensional matrix of shares Θ with typical element $\theta^h_i \geq 0$ which denotes the share of consumer h in the profit of the private firm i. Thus, we require $\sum_h \theta^h_i = 1$ for all $i \neq 0$. Let $\mathcal{O}$ denote the set of matrices Θ with these properties. A private ownership economy corresponding to a matrix of ownership shares $\Theta \in \mathcal{O}$ is characterized by $E(\Theta) = \langle (X^h, u^h), (Y^i), \Theta \rangle$.

3. A Heuristic Discussion of Earlier Strategies and Our Strategy.

Before launching into our analysis, it is first worth running over the central ideas in the proofs of Dasgupta and Stiglitz, Mirrlees, Hahn, Sadka, and Reinhorn. Dasgupta and Stiglitz assume that technologies of firms satisfy the Inada conditions. They employ differential techniques to prove that production efficiency is desirable at all second-best tax equilibria of private ownership economies with firm-specific profit taxation. With an aim at understanding the precise mechanism and the intuition behind this result, Mirrlees and Hahn present alternative proofs using non-differential techniques and for more general technologies.

The central idea behind these proofs is an application of an argument in the DM paper based on the assumption that there is always one commodity that every consumer likes and hence net demand is positive (or dislikes and hence net demand is negative).[4]

[3] For all h, X^h is net of the consumer h's endowment.
[4] For a more general condition that backs this argument, see Weymark [1978].

This assumption implies a more general assumption of local Pareto nonsatiation as defined by Hahn.[5]

To understand the intuition of the DM strategy, we consider the set of tax equilibria as defined by Model A below, where $m^h = m$ for all h. The DM model is a special case of Model A, where Y^i is a cone for all i. Guesnerie's model is also a special case of Model A where, for all i, Y^i exhibits decreasing returns to scale and there is one-hundred percent profit taxation of private firms.

Model A (DM/Guesnerie):

$$\sum_h x^h(q,m) \leq \sum_{i \neq 0} y^i(p) + y^0(p^0) \tag{3.1}$$

Walras Law implies that at a tax equilibrium of Model A, the government's budget is balanced:

$$Hm \leq [q-p] \sum_{i \neq 0} y^i(p) + qy^0(p^0) + p \sum_{i \neq 0} y^i(p), \tag{3.2}$$

that is, the indirect and profit tax revenues of the government are used to finance the demogrant expenditure of the government.

DM show that under local Pareto nonsatiation, the set of second-best Pareto optimal allocations of Model A is also the set of Pareto optimal allocations of a less constrained economy and the second-best Pareto optimal allocations of the less constrained economy are all production efficient. The equilibria of this less constrained economy are defined by Model B below, where the government is free to choose any production vector in Y and the consumers maximize utility.

Model B:

$$\sum_h x^h(q,m) \leq y + \sum_h e^h, \ y \in \sum_{i \neq 0} Y^i + Y^0. \tag{3.3}$$

Walras law implies that the government budget is balanced:

$$Hm \leq qy, \tag{3.4}$$

that is the government distributes the receipts from sale of y as demogrant incomes to consumers.

It is clear that the tax equilibrium allocations of Model A are a subset of the tax equilibrium allocations of Model B. If an equilibrium allocation of Model B is production inefficient, that is, lies in the interior of Y, then by DM version of local Pareto nonsatiation and continuity of net demands of consumers in consumer prices, a Pareto improving change in consumer prices that also leads to a technologically feasible change in aggregate demand

5 Local Pareto nonsatiation is also satisfied if the government is permitted to use a demogrant provided consumers' preferences are strictly monotonic.

always exists: a well-calibrated decrease (increase) in the consumer price of the good everyone likes (dislikes) ensures that the new aggregate demand lies in a small enough neighborhood of the original aggregate demand that is contained in Y. The new allocation is hence an equilibrium of Model B. This suggests that no production inefficient equilibrium of Model B can be a second-best of Model B as there always exists another equilibrium of Model B that Pareto dominates it. This implies that, at a second-best of Model B, there always exists a producer price vector that can be used to decentralize it as a (second-best) tax equilibrium of Model A.

In Model C below, the profits of the private firms are distributed to consumers according to an exogeneous matrix of shares $\Theta \in \mathcal{O}$ and every private firm i is subject to a firm-specific profit tax rate τ^i. The net income of consumer h is $m^h = m + \sum_{i \neq 0} \theta_i^h (1-\tau^i)\pi^i(p)$.

Model C (Private Ownership with Firm-Specific Profit taxation):

$$\sum_h x^h(q, \sum_{i \neq 0} \theta_i^h (1-\tau^i) p y^i(p) + m) \leq \sum_{i \neq 0} y^i(p) + y^0(p^0). \tag{3.5}$$

Walras Law implies that the government's budget is balanced at a tax equilibrium of Model C.[6]

$$Hm \leq [q-p] \sum_{i \neq 0} y^i(p) + q y^0(p^0) + p \sum_{i \neq 0} \tau^i y^i(p), \tag{3.6}$$

that is the government distributes its revenue from indirect and profit taxation as demogrant incomes to consumers.

With a view to understanding the intuition behind the Dasgupta and Stiglitz result, that is to check whether the second-best of Model C are production efficient, Mirrlees considers a more general model, Model C', where the private firms face firm-specific prices.

Model C' (Mirrlees):

$$\sum_h x^h(q, \sum_{i \neq 0} \theta_i^h p^i y^i(p^i) + R) \leq \sum_{i \neq 0} y^i(p^i) + y^0(p^0). \tag{3.7}$$

Mirrlees aims to find a less constrained model such that the Pareto optimal allocations of the less constrained model are all production efficient and correspond to the second-best allocations of Model C'. He considers Model D below that involves firm specific numbers r_i for all $i \neq 0$ and where the government can choose any production vector in Y.

Model D

$$\sum_h x^h(q, \sum_{i \neq 0} \theta_i^h r_i + R) \leq y, \ y \in \sum_{\neq 0} Y^i + Y^0. \tag{3.8}$$

[6] This is proved in Section 5.

Walras law implies that the government's budget is balanced:

$$HR + \sum_{i \neq 0} r_i \leq qy. \tag{3.9}$$

It is clear that any tax equilibrium allocation of Model C' can be obtained as an equilibrium allocation of Model D, where $r^i = p^i y^i(p^i)$ for all $i \neq 0$. Exactly as argued in the case of Model B, under the DM version of local Pareto nonsatiation, it can be shown that the Pareto optimal allocations of Model D are production efficient. The question is whether each Pareto optimal allocation of Model D can be decentralized as a tax equilibrium of Model C'? Production efficiency at the optimum of Model D implies that there exists a vector of producer prices (say p) that can decentralize the underlying production vectors as profit maximizing choices of the private firms. If the profits of all firms at this price vector are all non-zero, then Mirrlees shows that, for every $i \neq 0$, the profit of firm i can be suitably scaled (say by a factor λ^i) so that the scaled profit equals the value of the number r^i at the optimum (that is, $\lambda^i \pi^i(p) = r_i$). Note that, given the homogeneity properties of the profit and the supply functions, this is equivalent to firm i facing the firm specific price $p^i = \lambda^i p$ and maximizing profits (the supply vector of the firm remains unchanged, $y^i(p) = y^i(p^i)$, so that $\lambda^i \pi^i(p) = \pi^i(\lambda^i p) = r_i$). Thus, the given Pareto optimum of Model D is a tax equilibrium of Model C'. Moreover, it is also a tax equilibrium of Model C as these firm-specific scaling factors can be used to define the firm-specific profit tax rates by setting $(1 - \tau^i)$ equal to λ^i for all $i \neq 0$.[7]

However, this argument may fail if at a Pareto optimum of Model D there exists $i \neq 0$ such that r_i is not equal to zero but the profit of firm i, $\pi^i(p)$ is equal to 0. In that case there exists no scaling factor λ^i such that $\lambda^i \pi^i(p) = r_i$. Thus, for economies with more general technologies than in Dasgupta and Stiglitz, Mirrlees could not conclusively demonstrate production efficiency at a second-best of Model C.

Instead of searching for a less constrained problem whose Pareto optimal allocations are all production efficient and correspond to the second-best of Model C, Hahn applies the DM strategy based on the assumption of local Pareto nonsatiation directly to Model C.

[7] To be more accurate, second-best allocations of Model D, where and $\pi^i(p) \neq 0$ for all $i \neq 0$ can be decentralized as equilibria of Model C' only when $r^i \neq 0$ for all $i \neq 0$, for if $r^i = 0$ for some $i \neq 0$, then $\lambda^i = 0$ and the profit of firm i is not well-defined at the price $p^i = \lambda^i p = 0$. On the other hand such a the second-best allocation of Model D can always be decentralized as an equilibrium of Model C, for $\lambda^i = 0$ would imply that firm i faces price vector p but is subject to one-hundred percent profit taxation $(\tau^i = 1)$. This seems to be a common source of confusion in the earlier literature, which seems to suggest that an equilibrium allocation of Model C is always decentralizable as an equilibrium allocation of Model C'. This is not true when an equilibrium allocation of Model C involves a private firm that is subject to one hundred-percent profit taxation.

Note that in the DM proof of production efficiency at a second-best of Model A (which, as seen above, employs Model B), the move from a production inefficient equilibrium allocation to a Pareto superior equilibrium allocation of Model B is achieved without changing the incomes of the consumers (only consumer prices change to achieve the Pareto improvement). This means that the government's revenue (which is the receipts from sales to consumers and is equal to the aggregate net consumer expenditure) is unchanged in the move to the Pareto superior equilibrium. Thus, it can continue distributing the same demogrant, which is the sole component of the consumers' net incomes. Thus, the aggregate and individual net incomes/expenditures of consumers can be held constant in this Pareto improving move.

However, this strategy of proof does not work, in general, if we wish to exploit the assumption of local Pareto nonsatiation directly in Model C. Suppose a is a production inefficient tax equilibrium of Model C. If the net incomes of consumers can be held at the levels in a, then under the DM version of local Pareto nonsatiation and the continuity of net demands of consumers in consumer prices, there exists a sequence of Pareto dominating changes in consumer demands that converge to the demands at a such that, eventually, the aggregate demands become technologically feasible. At every point in this sequence, aggregate net expenditures and aggregate net incomes are the same as at a. To ensure that eventually all points in this sequence can be obtained as equilibria of Model C (that are Pareto superior to a) requires ensuring that the individual incomes along this sequence can indeed be maintained at the levels at a as hypothesized. However, the technologically feasible points of the sequence involve production levels that are different from a. Therefore the profits of individual firms, and hence profit incomes of consumers, may not be the same along this sequence as in a. Note that (3.7) implies that

$$\begin{aligned} HR \leq & q[\sum_{i\neq 0} y^i(p) + y^0(p^0)] - \sum_{i\neq 0}(1-\tau^i)py^i(p) \\ = & q[\sum_{i\neq 0} y^i(p) + y^0(p^0)] - \sum_{i\neq 0}\sum_h \theta_i^h(1-\tau^i)py^i(p), \end{aligned} \tag{3.10}$$

that is the amount distributed as demogrant is the aggregate income (total value of aggregate output at consumer prices) minus the aggregate net-of-tax profit incomes in the economy. Since aggregate net income is the same along this sequence, (3.10) implies that the level of the demogrant may also vary along this sequence. Thus, in general, the distribution of individual net incomes may differ from a along this sequence.

Nevertheless, Hahn attempts to show that, if technologies are smooth, then the profits associated with the technologically feasible points in the sequence converge to the profits at a. Hence there is a small neighborhood around a where the profits of individual firms associated with the sequence have the same sign as the profits at a (in particular, they are non-zero if profits at a are non-zero). By appropriately changing firm specific profit tax rates (scaling profits of individual firms), the profits of private firms, and hence the net profit incomes of consumers, can be held at the levels at a. Employing (3.10) and recalling that the aggregate net income is constant, this means the demogrant also takes the same value as at a. Hence, by implementing firm-specific profit taxation, eventually, the distribution of net incomes along the sequence can also be fixed at the levels at a. All such points are equilibria of Model C that are Pareto superior to a.

As pointed out by Reinhorn, the problem with Hahn's analysis is that, although the points on the sequence are eventually technologically feasible, it does not mean that the production vectors of individual firms along this sequence lie on the frontiers of their individual technologies. There may exist no changes in producer prices that induce profit maximizing firms to change their supplies to meet exactly the Pareto-improving change in the aggregate demand: this may happen when the Pareto-improving change in aggregate demand requires all firms to produce in the interiors of their individual technologies (as opposed to the frontiers of their technologies), which is never profit-maximizing.

Let the tax equilibrium configuration at a production inefficient tax equilibrium a of Model C be denoted by bars. Then $\sum_h x^h(\bar{q}, \bar{R}+\sum_{i\neq 0}\theta_i^h(1-\bar{\tau}^i)\pi^i(\bar{p})) = \bar{x} = \sum_{i\neq 0} y^i(\bar{p})+ y^0(\bar{p}^0)$. Since all firms maximize profits, it is clear that for all i, $\bar{y}^i$ lies on the frontier of Y^i. In his proof, Reinhorn, focuses on a subset (he denotes it as $K(\epsilon^1, \ldots, \epsilon^I)$) of the aggregate technology) obtained as $\sum_i [N_{\epsilon^i}(\bar{y}^i) \cap Y^i]$.[8] If the frontiers of individual technologies are smooth enough, then for all $i \neq 0$, there exist $\epsilon^i > 0$ such that the signs of profits of firm i at points that lie on the frontier of Y^i intersected with $N_{\epsilon^i}(\bar{y}^i) \cap Y^i$ are the same as at $\bar{y}^i$. Thus, as in Hahn and Mirrlees, such points can, through appropriate firm-specific profit taxation, yield the same net-of-tax profits to individual firms (and hence, the same profit incomes to consumers) as in the $\bar{x}$ situation. $\bar{x}$ clearly is in $K(\epsilon^1, \ldots, \epsilon^I)$. However, since $\bar{x}$ is not production efficient ($\bar{p}$ is not proportional to $\bar{p}^0$), it lies in the interior of the set $K(\epsilon^1, \ldots, \epsilon^I) - \mathbf{R}_+^N$, which is a subset of the aggregate technology (assuming free disposability of aggregate technology), so that the DM argument can be applied to isolate Pareto-improving changes in consumer prices. The problem with Reinhorn's argument (as in the case of Hahn) continues to be the fact that, while such changes can be calibrated to ensure that they lead to a change in aggregate demand that lies in the set $K(\epsilon^1, \ldots, \epsilon^I) - \mathbf{R}_+^N$, and hence is technologically feasible, they may not lead to points in the set $K(\epsilon^1, \ldots, \epsilon^I)$ which involve individual firms producing on the frontiers of their respective technologies, that is points that can be achieved by changing producer prices given profit maximizing behavior of producers.[9]

[8] For $y \in \mathbf{R}^n$, $N_\epsilon(y)$ is the ϵ neighborhood of y in $\mathbf{R}^n$.

[9] $K(\epsilon^1, \ldots, \epsilon^I) - \mathbf{R}_+^N$ includes points where individual firms may be producing in the interiors of their technologies.

We adopt a different proof strategy (see Theorem 1). Instead of first constructing Pareto-improving directions of change at a production inefficient status-quo (as the above papers did), we show first that at any production inefficient tax equilibria, there exist small-enough changes in the producer prices that (i) lead to an increase in the aggregate supply, and hence the aggregate income, in the economy (see Lemma 2 and its proof below) and (ii) maintain (with firm-specific taxation) the profits of firms, and hence profit incomes of consumers, for any given vector of profit shares, at the levels of the initial equilibrium. This implies that the increase in the aggregate income in the economy due to the change in producer prices shows up only as an increase in the tax revenue of the government, which (under the assumption of local Pareto nonsatiation) the government can use to change consumer prices (commodity taxes) or redistribute as uniform lump-sum transfers to increase welfare of all consumers in a non-tight equilibrium preserving way.

The same strategy is then used to demonstrate second-best production efficiency in some structures of private-ownership under strong institutional constraints on profit taxation, that is, when there can be no profit taxation (or there is, at best, uniform profit taxation).

4. Some Preliminary Lemmas.

Assumptions 1, 2 and 3 stated below are further regularity assumptions on the technologies of firms.

Assumption 1. For all i, Y^i is closed, not a cone, and satisfies free-disposability ($Y^i + \mathbf{R}^N_- \subset Y^i$).

Assumption 2. For all i, the set B^i is non-empty and there exists $\rho \in \mathbf{R}^N_{++} \cap B^i$.

Define the frontier of Y as

$$\hat{Y} = \{y \in Y | \bar{y} \gg y \text{ implies that } \bar{y} \notin Y\}. \tag{4.1}$$

Similarly, we can define $\hat{Y}^i$ for all i, $\hat{Y}^c$, and $\hat{Y}^0$.[10] Note that Assumption 1 guarantees that the frontier of Y and the boundary of Y (denoted by ∂Y) coincide. Further note that Assumptions 1 and 2 imply that for any i, if $\bar{y}$ lies in Y^i, then there exist production points in $\hat{Y}^i$ that are no smaller than $\bar{y}$. In particular, if $\bar{y}$ lies in the interior of Y^i, then there are production points in in $\hat{Y}^i$ that are strictly larger than $\bar{y}$.

[10] Vector notation: for x and y in $\mathbf{R}^n$,

$$\begin{aligned} &x \geq y \Leftrightarrow x_i \geq y_i, \ \forall i = 1, \dots, n, \\ &x > y \Leftrightarrow x \neq y \text{ and } x_i \geq y_i, \ \forall i = 1, \dots, n, \\ &x \gg y \Leftrightarrow x_i > y_i, \ \forall i = 1, \dots, n, \end{aligned} \tag{4.2}$$

.

Assumption 1 also excludes firms that exhibit constant returns-to-scale. This exclusion seems without loss of generality as, firstly, it is meaningless to assume that consumers can have shares in the profits of such firms and, secondly, the presence of such firms offers no constraints on distribution of profits in the economy (the issue of focus in this paper). Note that under Assumptions 1 and 2, for all i, π^i is continuous, linearly homogeneous, convex on the set B^i. We denote the aggregate supply of the private competitive sector by $y^c(\rho) = \sum_{i \neq 0} y^i(\rho)$.[11]

Assumption 3. There exist smooth and strictly quasi-convex functions $f^c : \mathbf{R}^N \to \mathbf{R}$ and $f^0 : \mathbf{R}^N \to \mathbf{R}$ with images $f^c(y)$ and $f^0(y)$ such that $Y^c = \{y \in \mathbf{R}^N | f^c(y) \leq 0\}$ and $Y^0 = \{y \in \mathbf{R}^N | f^0(y) \leq 0\}$.[12]

If both Assumptions 1 and 3 hold, then $\hat{Y}^c = \{y \in \mathbf{R}^N | f^c(y^c) = 0\}$ and $\hat{Y}^0 = \{y \in \mathbf{R}^N | f^0(y^0) = 0\}$.

Lemmas 1, 2, and the Corollary to Lemma 2 demonstrate that a vector of aggregate supply is production efficient iff the price vectors faced by the private sector firms and government sector firm are proportional, and if the firms face different (non-proportional) producer prices, then there exist changes in producer prices of the private and government firms that can strictly increase the aggregate supply. Intuitively, the proof of Lemma 2 exploits the differences in the marginal rates of substitution in production that exist when the price vectors of all firms are not proportional to construct such changes in prices.

Lemma 1: *Let Assumptions 1 and 2 hold and let $p \in B^c$ and $p^0 \in B^0$. Then $y^c(p) + y^0(p^0) \notin \hat{Y}$ implies that p is not proportional to p^0.*

Lemma 2: *Suppose Assumptions 1, 2, and 3 hold. Let $\bar{p} \in B^c$ and $\bar{p}^0 \in B^0$ such that $\bar{p}$ and $\bar{p}^0$ are not proportional. Let $\bar{y} := y^c(\bar{p}) + y^0(\bar{p}^0)$. Then there exist sequences $\{p^v\} \to \bar{p}$ and $\{p^{0^v}\} \to \bar{p}^0$ such that $y^c(p^v) + y^0(p^{0^v}) \gg \{\bar{y}\}$ for all large enough v.*

Corollary to Lemma 2: *Suppose Assumptions 1 to 3 hold. Let $\bar{p} \in B^c$ and $\bar{p}^0 \in B^0$ such that $\bar{y} = y^c(\bar{p}) + y^0(\bar{p}^0)$. Then $\bar{y} \notin \hat{Y}$ if $\bar{p}$ and $\bar{p}^0$ are not proportional.*

[11] The domain of y^c is $\cap_{i \neq 0} B^i = B^c$.

[12] The restriction to strictly-quasi convex functions f^c and f^0 implies that the supply mappings of the private and the government sector firms are functions. This is purely for convenience. Theorem 1 can be generalized to the case where these functions are quasi-convex.

5. Firm-Specific Profit Taxation and Production Efficiency.

In this section we show that all the second-best allocations in such an economy are production efficient. For every $i \neq 0$, the profit tax rate is τ^i, where $\tau^i \leq 1$.

Definition. For any $\Theta \in \mathcal{O}$, a tax equilibrium of a private ownership economy $E(\Theta)$ with firm-specific profit taxes for private firms is a configuration $\langle q, p^0, p, \tau^1, \ldots, \tau^I, m\rangle \in \mathbf{R}^N_{++} \times B^0 \times B^c \times \mathbf{R}^I \times \mathbf{R}$ such that (3.7) holds.

For every $\Theta \in \mathcal{O}$, let $\mathcal{E}_\tau(\Theta)$ be the set of tax-equilibrium configurations of q, p, p^0, $\tau^1, \ldots, \tau^I$, and m in the economy $E(\Theta)$ with firm-specific profit taxation. The system of equations (3.7) is homogeneous of degree zero in p^0. It is also homogeneous of degree zero in p, q, and m. Hence, it admits two normalizations. We adopt the normalization rules $p^0_1 = 1$ and and $p_1 = 1$.

Under the maintained assumptions on consumers' preferences, the budget constraints hold as equalities under utility maximization, that is, for all h, we have

$$q \cdot x^h(q, \sum_{i\neq 0} \theta^h_i (1-\tau^i)\pi^i(p) + m) = \sum_{i\neq 0}(1-\tau^i)\theta^h_i \pi^i(p) + m. \tag{5.1}$$

To show that if $e \in \mathcal{E}_\tau(\Theta)$ then the government budget is balanced, we multiply both sides of (3.7) by q and employ (5.1) to obtain

$$\begin{aligned}
q^T \sum_h x^h(q, m^h) &\leq q^T \sum_{i\neq 0} y^i(p) + q^T y^0(p^0) \\
\Rightarrow \sum_h [\sum_{i\neq 0} \theta^h_i (1-\tau^i)\pi^i(p) + m] &\leq \sum_{i\neq 0} p^T y^i(p) + \sum_{i\neq 0}[q^T - p^T] y^i(p^i) + q^T y^0(p^0) \\
\Rightarrow Hm &\leq q^T y^0(p^0) + \sum_{i\neq 0}[q^T - p^T] y^i(p) + \sum_{i \neq 0} \tau^i p y^i(p) \\
\Rightarrow m &\leq \frac{q^T y^0(p^0) + \sum_{i\neq 0}[q^T - p^T] y^i(p) + \sum_{i\neq 0} \tau^i p y^i(p)}{H}.
\end{aligned} \tag{5.2}$$

Condition (5.2) says that the demogrant is financed from the government's revenue from indirect taxation ($\sum_{i\neq 0}[q^T - p^T]y^i(p)$), profit taxation ($\sum_{i\neq 0} \tau^i p y^i(p)$), and receipts from sale of publically produced private commodities ($q^T y^0(p^0)$).

The following theorem establishes that every second-best allocation of a private ownership economy with firm-specific profit taxation is production efficient.

The second-best problem is to find the mapping $\mathcal{V}_\tau : \Delta_{H-1} \times \mathcal{O} \to \mathbf{R}$ with image[13]

$$\mathcal{V}_\tau(s^1, \ldots, s^H, \Theta) :=$$

$$\max_{q\in\mathbf{R}^N_{++},\ (\tau^i)_{i\neq 0}\in\mathbf{R}^I,\ m\in\mathbf{R},\ p\in\mathbf{R}^N_+,\ p^0\in\mathbf{R}^N_+} \sum_h s_h u^h(x^h(q, \sum_{i\neq 0} \theta^h_i(1-\tau^i)\pi^i(p) + m)) \qquad (*)$$

$$\text{subject to}$$

$$\langle q, p, p^0, (\tau^i)_{i\neq 0}, m\rangle \in \mathcal{E}_\tau(\Theta).$$

Theorem 1: *Suppose Assumptions 1, 2, and 3 hold. Suppose either*[14]
(a) there exists h such that u^h is strictly monotonic or
(b) there exists a commodity k such that either of the following holds:
(i) $\sum_h x^h_k(q, m^h) > 0,\ x^h_k(q, m^h)\sum_h x^h_k(q, m^h) \geq 0,\ \forall h,\ q \in \mathbf{R}^N_{++},$ and $m^h \in \mathbf{R}$ or
(ii) $\sum_h x^h_k(q, m^h) < 0,\ x^h_k(q, m^h)\sum_h x^h_k(q, m^h) \geq 0,\ \forall h,\ q \in \mathbf{R}^N_{++},$ and $m^h \in \mathbf{R}$
If $\bar{e} := \langle \bar{q}, \bar{p}, \bar{p}^0, (\bar{\tau}^i)_{i\neq 0}, \bar{m}\rangle \in \mathbf{R}^{3N+I+1}$ is a solution to $(*)$ then $y^c(\bar{p}) + y^0(\bar{p}^0) \in \hat{Y}$.

Proof: Suppose $\bar{y} := y^c(\bar{p}) + y^0(\bar{p}^0) \notin \hat{Y}$. From Lemma 1, Lemma 3, and corollary to Lemma 2 this implies that $\bar{y} \in$ Interior Y and $\bar{p}^0 \neq \kappa\bar{p}$ for any $\kappa \geq 0$.
Step 1. We show that there exist changes in the producer prices which leads to aggregate supplies that are greater than $\bar{y}$. This is true because Lemma 2 demonstrates that there exist sequences $\{p^v\} \to \bar{p}$ and $\{p^{0^v}\} \to \bar{p}^0$ such that $y^c(p^v) + y^0(p^{0^v}) \gg \bar{y}$ for all $v > \hat{v}$, where $\hat{v}$ is defined as in the proof of Lemma 2.

This implies that the aggregate income or the value of aggregate output measured using consumer prices $\bar{q}$ increases when producer prices $\bar{p}$ and $\bar{p}^0$ change to p^v and p^{0^v} for all $v > \hat{v}$, that is,

$$M^v = \bar{q}^T[y^c(p^v) + y^0(p^{0^v})] > \bar{q}^T\bar{y} =: \bar{M},\ \forall v > \hat{v}. \qquad (5.3)$$

Step 2. We show that there exists a $\overset{*}{v} > \hat{v}$ and firm specific profit taxes $\overset{*}{\tau}{}^i$ for all $i \neq 0$ such that

$$(1 - \overset{*}{\tau}{}^i)\pi^i(p^{\overset{*}{v}}) = (1 - \bar{\tau}^i)\pi^i(\bar{p}),\ \forall i \neq 0. \qquad (5.4)$$

The continuity of the profit functions π^i implies that $\pi^i(p^v) \to \pi^i(\bar{p})$. Let $\mathcal{I} = \{i \neq 0 |\ (1 - \bar{\tau}^i)\pi^i(\bar{p}) > 0\}$. If $\mathcal{I} = \emptyset$, then choose any $\epsilon > 0$. If $\mathcal{I} \neq \emptyset$, then choose ϵ such that $0 < \epsilon < \min_{i\in\mathcal{I}}\{(1 - \bar{\tau}^i)\pi^i(\bar{p})\}$. For every $i \neq 0$ there exists v^i such that for all $v > v^i$ we have $|\pi^i(p^v) - (1 - \bar{\tau}^i)\pi^i(\bar{p})| < \epsilon$. Our choice of ϵ implies that, for every $i \neq 0$ and $v > v^i$, the sign of $\pi^i(p^v)$ is the same as the sign of $(1 - \bar{\tau}^i)\pi^i(\bar{p})$: if $(1 - \bar{\tau}^i)\pi^i(\bar{p}) = 0$ then $\pi^i(p^v) \geq 0$ and if $(1 - \bar{\tau}^i)\pi^i(\bar{p}) > 0$ then $\pi^i(p^v) > 0$.

[13] Δ_{H-1} is the $H-1$-dimensional unit simplex in $\mathbf{R}^H$.
[14] Note that since the government can employ a demogrant, (a) below ensures that local Pareto nonsatiation and (b) below is the DM version of local Pareto nonsatiation.

Pick $\overset{*}{v}$ to be any $v > \max\{v^1, \dots, v^I, \hat{v}\}$. Define $p^{\overset{*}{v}} =: \overset{*}{p}$, $p^{0^{\overset{*}{v}}} := \overset{*}{p}^0$, and $\overset{*}{y}^c := y^c(\overset{*}{p})$. For $i \neq 0$, choose firm specific scaling factors λ^i such that

$$\lambda^i \pi^i(\overset{*}{p}) = (1 - \bar{\tau}^i)\pi^i(\bar{p}). \tag{5.5}$$

This is possible, for if $\mathcal{I} = \emptyset$, then $\lambda^i = 0$ needs to be chosen for all $i \neq 0$. If $\mathcal{I} \neq \emptyset$, then for all $i \neq 0$ and and not in $\mathcal{I}$, choose $\lambda^i = 0$ and for $i \in \mathcal{I}$, λ^i is given by

$$\lambda^i = \frac{(1 - \bar{\tau}^i)\pi^i(\bar{p})}{\pi^i(\overset{*}{p})}, \tag{5.6}$$

which is well defined as $\pi^i(\overset{*}{p}) \neq 0$ for $i \in \mathcal{I}$. Note, for all $i \neq 0$, $\lambda^i \geq 0$.

Define $\overset{*}{\tau}^i = 1 - \lambda^i$ for all $i \neq 0$. Then $1 - \overset{*}{\tau}^i \geq 0$, which implies, $\overset{*}{\tau}^i \leq 1$. Then (5.6) implies (5.4) and

$$\sum_{i \neq 0} \theta_i^h (1 - \bar{\tau}^i)\pi^i(\bar{p}) = \sum_{i \neq 0} \theta_i^h (1 - \overset{*}{\tau}^i)\pi^i(\overset{*}{p}^i), \ \forall h. \tag{5.7}$$

Summing up over all h, this implies that

$$\sum_{i \neq 0} (1 - \overset{*}{\tau}^i)\pi^i(\overset{*}{p}) = \sum_{i \neq 0} (1 - \bar{\tau}^i)\pi^i(\bar{p}). \tag{5.8}$$

For all $i \neq 0$, $\overset{*}{\tau}^i$ can be interpreted as the firm-specific profit tax rate. Thus, the individual and the aggregate net-of-tax profits of firms and consumers' profit incomes does not change when $\bar{p}$ changes to $\overset{*}{p}$ and when the profit tax rates $\overset{*}{\tau}^i$ apply for all $i \neq 0$.

Step 3. Since the move to $\overset{*}{p}$, $\overset{*}{p}^0$, and $\overset{*}{\tau}^i$ for all $i \neq 0$ results in no change in the net-of-tax profit incomes of consumers, it must imply that the increase in aggregate income from $\bar{M}$ to $M^{\overset{*}{v}}$ must show up as an increase in the government's tax revenue and its public sector activities. This is what we show now.

$$\begin{aligned}
M^{\overset{*}{v}} =: \overset{*}{M} &= \bar{q}^T [\sum_{i \neq 0} y^i(\overset{*}{p}) + y^0(\overset{*}{p}^0)] \\
&= [\bar{q}^T - \overset{*}{p}] \sum_{i \neq 0} y^i(\overset{*}{p}) + \overset{*}{p} \sum_{i \neq 0} y^i(\overset{*}{p}) + \bar{q}^T y^0(\overset{*}{p}^0) \\
&= [\bar{q}^T - \overset{*}{p}] \sum_{i \neq 0} y^i(\overset{*}{p}) + \sum_{i \neq 0} \overset{*}{\tau}^i \overset{*}{p} y^i(\overset{*}{p}) + \sum_{i \neq 0} (1 - \overset{*}{\tau}^i) y^i(\overset{*}{p}) + \bar{q}^T y^0(\overset{*}{p}^0) \\
&= [\bar{q}^T - \overset{*}{p}] \sum_{i \neq 0} y^i(\overset{*}{p}) + \sum_{i \neq 0} \overset{*}{\tau}^i \pi^i(\overset{*}{p}) + \sum_{i \neq 0} (1 - \overset{*}{\tau}^i)\pi^i(\overset{*}{p}) + \bar{q}^T y^0(\overset{*}{p}^0).
\end{aligned} \tag{5.9}$$

Similarly,

$$\bar{M} = [\bar{q}^T - \bar{p}] \sum_{i \neq 0} y^i(\bar{p}) + \sum_{i \neq 0} \bar{\tau}^i \pi^i(\bar{p}) + \sum_{i \neq 0} (1 - \bar{\tau}^i)\pi^i(\bar{p}) + \bar{q}^T y^0(\bar{p}^0). \tag{5.10}$$

Since $\overset{*}{M} - \bar{M} > 0$, it follows from (5.8) that the government's revenue from commodity and profit taxes and its public sector activities is higher when we move to $\overset{*}{p}$, $\overset{*}{p}{}^0$, and $\overset{*}{\tau}{}^i$ for all $i \neq 0$, keeping consumer prices unchanged, that is,

$$\begin{aligned} \overset{*}{G} :=& [\bar{q}^T - \overset{*}{p}] \sum_{i \neq 0} y^i(\overset{*}{p}) + \sum_{i \neq 0} \overset{*}{\tau}{}^i \pi^i(\overset{*}{p}) + \bar{q}^T y^0(\overset{*}{p}{}^0) \\ &> [\bar{q}^T - \bar{p}] \sum_{i \neq 0} y^i(\bar{p}) + \sum_{i \neq 0} \bar{\tau}^i \pi^i(\bar{p}) + \bar{q}^T y^0(\bar{p}^0) =: \bar{G}. \end{aligned} \tag{5.11}$$

Step 4. Next we show that this increase in the government's revenue can be used to construct another tax equilibrium where utility of at least one consumer is higher, with no loss in utility for the others: this is obtained by either increasing the demogrant by an appropriate amount (this is possible if (a) holds) or by decreasing (increasing) the consumer price of, and hence the tax on, the k$^{\text{th}}$ commodity by an appropriate amount (this is possible if b(i) (b(ii)) holds).

For all h define $x^h(\bar{q}, \sum_{i \neq 0} \theta_i^h (1 - \bar{\tau}^i) \pi^i(\bar{p}) + \bar{m}) =: \bar{x}^h$ and $\sum_h \bar{x}^h =: \bar{x}$. (5.7) implies that for all h, $\bar{x}^h = x^h(\bar{q}, \sum_{i \neq 0} \theta_i^h (1 - \overset{*}{\tau}{}^i) \pi^i(\overset{*}{p}) + \bar{m}))$. Since $\bar{x} \leq \bar{y} \ll y^c(\overset{*}{p}) + y^0(\overset{*}{p}{}^0) =: \overset{*}{y}$, we have $\bar{x} \in$ Interior $\{\overset{*}{y}\} + \mathbf{R}^N_{--}$.

Since x^h is a continuous function of q_k for all h, clearly, if condition b(i) or b(ii) hold, we can apply the DM argument to find $\overset{*}{q}_k$ and $\overset{*}{\epsilon}$ such that (1) $\sum_h x^h(\bar{q}_{-k}, \overset{*}{q}_k, \sum_{i \neq 0} \theta_i^h (1 - \overset{*}{\tau}{}^i) \pi^i(\overset{*}{p}) + \bar{m})) =: \sum_h \overset{*}{x}{}^h \in N_{\overset{*}{\epsilon}}(\bar{x}) \subset \{\overset{*}{y}\} + \mathbf{R}^N_{--}$ and (2) $u^h(\overset{*}{x}{}^h) \geq u^h(\bar{x}^h)$ for all h and $u^h(\overset{*}{x}{}^h) > u^h(\bar{x}^h)$ for some h.

This implies that $\overset{*}{e} := \langle (\bar{q}_{-k}, \overset{*}{q}_k), \overset{*}{p}, (\overset{*}{\tau}{}^i)_{i \neq 0}, \overset{*}{p}{}^0 \rangle$ is another tax equilibrium configuration of $E(\Theta)$ that Pareto dominates $\bar{e}$.[15] To make the new tax equilibrium configuration conform to the normalization rules adopted, we exploit the homogeneity of the tax system to construct another tax equilibrium that is equivalent to $\overset{*}{e}$ and that satisfies the normalization rule, namely, $\langle (\bar{q}_{-k}, \overset{*}{q}_k), \frac{1}{\overset{*}{p}_1} \overset{*}{p}, (\overset{*}{\tau}{}^i)_{i \neq 0}, \frac{1}{\overset{*}{p}{}^0_1} \overset{*}{p}{}^0 \rangle$. This contradicts the fact that $\bar{e}$ is a solution to $(*)$.

If condition (a) holds, then we can exploit the continuity of x^h in m for all h to find $\hat{m} > \bar{m}$ and $\hat{\epsilon}$ such that (1) $\sum_h x^h(\bar{q}, \sum_{i \neq 0} \theta_i^h (1 - \overset{*}{\tau}{}^i) \pi^i(\overset{*}{p}) + \hat{m})) =: \sum_h \hat{x}^h \in N_{\overset{*}{\epsilon}}(\bar{x}) \subset \{\overset{*}{y}\} + \mathbf{R}^N_{--}$ and (2) $u^h(\hat{x}^h) \geq u^h(\bar{x}^h)$ for all h and $u^h(\overset{*}{x}{}^h) > u^h(\bar{x}^h)$ for some h.[16] This again leads to a tax equilibrium (that can be appropriately normalized) that Pareto dominates $\bar{e}$, which contradicts the hypothesis of the theorem. ■

[15] Note, this is made possible by the fact that $\overset{*}{G} > \bar{G}$ in (5.11), so that it is possible to use all or a part of this increased government budget-surplus to reduce taxes on some commodities. Also, note that $\overset{*}{e}$ is a non-tight tax equilibrium.

[16] Note, this is made possible by the fact that $\overset{*}{G} > \bar{G}$ in (5.11), so that it is possible to distribute all or a part of this increased government budget-surplus as a higher demogrant.

6. Some Cases of Second-Best Production Efficiency With Constraints on Profit Taxation.

In this section, we use a mechanism similar to the one in Theorem 1 to identify certain cases of private ownership where production efficiency holds at any second-best with restrictions on the government's ability to implement profit taxation. Specifically, we assume that the government can implement only uniform profit taxation.[17] These cases involve a rank conditions either on the matrix of ownership shares or on the matrix of supplies by the private producers, or more generally, on the product of these two matrices. This rank condition is made precise in Lemma 5 below. One class of economies that satisfies it occurs when $\theta_i^h = \theta^h$ for all h. The case where $\theta_i^h = \frac{1}{H}$ is a special case of this class and is equivalent to the case where government can tax away all profits of private firms at 100-percent and redistribute the tax revenue to consumers as a demogrant.

In this section, we make stronger assumptions on the technologies of the private firms: we require the technologies to satisfy the Inada conditions.

Assumption 4. For all $i \neq 0$, there exist smooth and strictly quasi-convex functions $f^i : \mathbf{R}^N \to \mathbf{R}$ with images $f^i(y)$, such that $Y^i := \{y \in \mathbf{R}^N | f^i(y) \leq 0\}$ includes zero and $Y^i \cap (-Y^i) = \emptyset$. For all $i \neq 0$, $B^i = \mathbf{R}^N_{++}$.

Remark 1. Assumptions 1 and 4 imply that for all $i \neq 0$, π^i is differentiable on $\mathbf{R}^N_{++}$, $\pi^i(p) > 0$ for all $p \in \mathbf{R}^N_{++}$, and $y^i(p) \neq 0^N$ for $p \gg 0^N$.[18]

The gross profit incomes of all consumers corresponding to producer price vector p can be expressed in the following matrix notations:

$$\begin{aligned}
\begin{bmatrix} \sum_{i\neq 0} \theta_i^1 \pi^i(p) \\ \vdots \\ \sum_{i\neq 0} \theta_i^H \pi^i(p) \end{bmatrix} &= \begin{bmatrix} \theta_1^1 & \dots & \theta_I^1 \\ \vdots & \vdots & \vdots \\ \theta_1^H & \dots & \theta_I^H \end{bmatrix} \begin{bmatrix} y_1^1(p) & \dots & y_N^1(p) \\ \vdots & \vdots & \vdots \\ y_1^I(p) & \dots & y_N^I(p) \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i\neq 0} \theta_i^1 y_1^i(p) & \dots & \sum_{i\neq 0} \theta_i^1 y_N^i(p) \\ \vdots & \vdots & \vdots \\ \sum_{i\neq 0} \theta_i^H y_1^i(p) & \dots & \sum_{i\neq 0} \theta_i^H y_N^i(p) \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i\neq 0} \theta_i^1 y^{iT}(p) \\ \vdots \\ \sum_{i\neq 0} \theta_i^H y^{iT}(p) \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} \\
&=: \Theta\Delta(p)p.
\end{aligned} \tag{6.1}$$

[17] Zero and hundred-percent profit taxation are special cases.

[18] This is because Assumptions 1 and 4 imply that, for all $i \neq 0$, $y^i \to 0^N$ implies there exists k such that $\nabla_{y_k^i} f^i(y^i) \to 0$.

Lemma 3 says that if the rows of the matrix $\Theta\Delta(p)$ span a one dimensional vector space for all $p \gg 0^N$, then under our assumptions on the technologies, each row is a constant (independent of p) linear multiple of the other.

Lemma 3: *Suppose Assumptions 1 and 4 hold, $\Theta \in \Omega$ is such that $\theta_i^h > 0$ for all $i \neq 0$ and for all h, and the rank of the matrix $\Theta\Delta(p)$ is one. Pick any $h' \in \{1, \dots, H\}$. Then there exist non-negative scalars μ^h for all h such that for all $p \in \mathbf{R}^N_{++}$, we have*

$$\sum_{i \neq 0} \theta_i^h y^{iT}(p) = \mu^h \sum_{i \neq 0} \theta_i^{h'} y^{iT}(p). \tag{6.2}$$

Remark 2. If $\theta_i^h = \theta^h$ for all h and $i \neq 0$ then rank of the matrix $\Theta\Delta(p)$ is one. If $\theta_i^h = \frac{1}{H}$ then rank of the matrix $\Theta\Delta(p)$ is one.

Remark 3. The rank of $\Theta\Delta(p)$ is one if the rank of Θ is one or rank of $\Delta(p)$ is one. The rank of Θ is one if and only if $\theta_i^h = \theta^h$ for all h and $i \neq 0$.

Definition. For any $\Theta \in \mathcal{O}$, a tax equilibrium of a private ownership economy $E(\Theta)$ with uniform profit taxation is a configuration $\langle q, p^0, p, m, \tau\rangle \in \mathbf{R}^N_{++} \times B^0 \times \mathbf{R}^N_{++} \times \mathbf{R} \times (-\infty, 1]$ such that

$$\sum_h x^h(q, \sum_{i \neq 0} \theta_i^h (1-\tau)\pi^i(p) + m) \leq \sum_{i \neq 0} y^i(p) + y^0(p^0). \tag{6.3}$$

For every $\Theta \in \mathcal{O}$, let $\mathcal{E}(\Theta)$ be the set of tax-equilibrium configurations of q, p, p^0, and m in the economy $E(\Theta)$. The system of equations (6.3) is homogeneous of degree zero in p^0. It is also homogeneous of degree zero in p, q, τ, and m. Hence, it admits two normalizations. We adopt the normalization rules $p_1^0 = 1$ and and $p_1 = 1$.

The second-best problem is to find the mapping $\mathcal{V} : \Delta_{H-1} \times \mathcal{O} \to \mathbf{R}$ with image

$$\mathcal{V}(s^1, \dots, s^H, \Theta) :=$$

$$\max_{q \in \mathbf{R}^N_{++},\ m \in \mathbf{R},\ p \in \mathbf{R}^N_{++},\ p^0 \in \mathbf{R}^N_{+}, \tau \in (-\infty, 1]} \sum_h s_h u^h(x^h(q, \sum_{i \neq 0} \theta_i^h (1-\tau)\pi^i(p) + m)) \tag{$**$}$$

subject to

$$\langle q, p, p^0, m, \tau\rangle \in \mathcal{E}(\Theta).$$

Theorem 2: *Suppose Assumptions 1 to 4 hold, $\Theta \in \Omega$ is such that $\theta_i^h > 0$ for all $i \neq 0$ and for all h, and the rank of the matrix $\Theta\Delta(p)$ is one. Suppose either*

(a) there exists h such that u^h is strictly monotonic or

(b) there exists a commodity k such that either of the following holds:

(i) $\sum_h x_k^h(q, m^h) > 0,\ x_k^h(q, m^h) \sum_h x_k^h(q, m^h) \geq 0,\ \forall h,\ q \in \mathbf{R}^N_{++},$ and $m^h \in \mathbf{R}$ or

(ii) $\sum_h x_k^h(q, m^h) < 0,\ x_k^h(q, m^h) \sum_h x_k^h(q, m^h) \geq 0,\ \forall h,\ q \in \mathbf{R}^N_{++},$ and $m^h \in \mathbf{R}$.

If $\bar{e} := \langle \bar{q}, \bar{p}, \bar{p}^0, \bar{m}, \bar{\tau}\rangle \in \mathbf{R}^{3N+I+1}$ is a solution to $(**)$ then $y^c(\bar{p}) + y^0(\bar{p}^0) \in \hat{Y}$.

Proof: Suppose $\bar{y} := y^c(\bar{p}) + y^0(\bar{p}^0) \notin \hat{Y}$. From Lemma 1, Lemma 3, and Corollary to Lemma 2 this implies that $\bar{y} \in$ Interior Y and $\bar{p}^0 \neq \kappa\bar{p}$ for any $\kappa \geq 0$.

Step 1. As in the proof of Theorem 1.

Step 2. We show that there exists a $\overset{*}{v} > \hat{v}$ and a common scaling factor $\overset{*}{\lambda} > 0$ such that

$$\sum_{i\neq 0} \theta_i^h \overset{*}{\lambda}\pi^i(p^{\overset{*}{v}}) = \mu^h \sum_{i\neq 0} \theta_i^{h'}(1-\bar{\tau})\pi^i(\bar{p}), \ \forall h. \tag{6.4}$$

The continuity of the profit functions π^i implies that $\pi^i(p^v) \to \pi^i(\bar{p})$. Remark 1 implies that, for all $i \neq 0$, $\pi^i(p^v) > 0$ for all v and $\pi^i(\bar{p}) > 0$.

Pick $\overset{*}{v}$ to be any $v > \hat{v}$. Define $p^{\overset{*}{v}} =: \overset{*}{p}$, $p^{0^{\overset{*}{v}}} := \overset{*}{p}^0$, and $\overset{*}{y}^c := y^c(\overset{*}{p})$. It is clear that $\pi^i(\overset{*}{p}) > 0$ for all $i \neq 0$. Pick any h' as in Lemma 3, and choose $\overset{*}{\lambda}$ such that

$$\sum_{i\neq 0} \theta_i^{h'}\pi^i(\overset{*}{\lambda}\overset{*}{p}) = \sum_{i\neq 0} \theta_i^{h'}(1-\bar{\tau})\pi^i(\bar{p}). \tag{6.5}$$

Exploiting the linear homogeneity of the profit function, $\overset{*}{\lambda}$ is given by

$$\overset{*}{\lambda} = \frac{\sum_{i\neq 0}\theta_i^{h'}\pi^i(\bar{p})}{\sum_{i\neq 0}\theta_i^{h'}(1-\bar{\tau})\pi^i(\overset{*}{p})}. \tag{6.6}$$

$\overset{*}{\lambda}$ is well defined, as the denominator in (6.6) is not zero. From Lemma 3, we know that there exist non-negative scalars μ^h for all h such that for all $p \in \mathbf{R}^N_{++}$, we have

$$\sum_{i\neq 0}\theta_i^h y^{iT}(p) = \mu^h \sum_{i\neq 0}\theta_i^{h'} y^{iT}(p). \tag{6.7}$$

From (6.5) and (6.7) it follows that for all h, we have

$$\sum_{i\neq 0}\theta_i^h \pi^i(\overset{*}{\lambda}\overset{*}{p}) = \mu^h \sum_{i\neq 0}\theta_i^{h'}\pi^i(\overset{*}{\lambda}\overset{*}{p}) = \mu^h \sum_{i\neq 0}\theta_i^{h'}(1-\bar{\tau})\pi^i(\bar{p}). \tag{6.8}$$

Step 3. Since the move to $\overset{*}{\lambda}\overset{*}{p}$ and $\overset{*}{p}^0$ results in no change in the net-of-tax profit incomes of consumers, it must imply that the increase in aggregate income from $\bar{M}$ to $M^{\overset{*}{v}}$ must show up as an increase in the government's tax revenue and its public sector activities. As in Steps 3 and 4 of the proof of Theorem 1, it can be shown that this increase in the government's revenue can be used to construct a tax equilibrium (with the adopted price normalizations) that Pareto dominates $\bar{e}$. ■

7. An Example of Production Inefficiency.

Consider the case when N is even, there are $\frac{N}{2}$ inputs, an equal number of outputs, and an equal number of competitive firms, that is, $I = \frac{N}{2}$. Each firm uses one input to produce one output, and the input and output are distinct for each firm. Suppose $H \leq \frac{N}{2}$ and Assumption 4 holds. This implies that $\Delta(p)$, given by

$$\begin{bmatrix} y_1^1(p) & 0 & \dots & 0 & y^1_{\frac{N}{2}+1}(p) & 0 & \dots & 0 \\ 0 & y_2^2(p) & \dots & 0 & 0 & y^2_{\frac{N}{2}+2}(p) & \dots & 0 \\ \vdots & \vdots & \vdots & & & & & \\ 0 & 0 & \dots & y^I_{\frac{N}{2}}(p) & 0 & 0 & \dots & y^I_N(p) \end{bmatrix}, \tag{7.1}$$

is full row-ranked. Suppose Θ is a full row-ranked matrix (an example is one where $H = I = \frac{N}{2}$ and Θ is an identity matrix). Then $\Theta\Delta(p)$ is full-row ranked. Suppose the government can implement uniform profit taxation. Using the indirect utility functions of consumers $(V^h(q, \sum_{i\neq 0} \theta_i^h(1-\tau)\pi^i(p) + m) = u^h(x^h(q, \sum_{i\neq 0} \theta_i^h(1-\tau)\pi^i(p) + m)))$, the second-best problem $(**)$ becomes

$$\mathcal{V}(s^1, \dots, s^H, \Theta) :=$$
$$\max_{q\in\mathbf{R}^N_{++},\ m\in\mathbf{R},\ p\in\mathbf{R}^N_{+},\ p^0\in\mathbf{R}^N_{+}, \tau\in(-\infty,1]} \sum_h s_h V^h(q, \sum_{i\neq 0} \theta_i^h(1-\tau)\pi^i(p) + m) \tag{$***$}$$
$$\text{subject to}$$
$$\langle q, p, p^0, m, \tau\rangle \in \mathcal{E}(\Theta).$$

The Lagrangian of the problem is

$$L =$$
$$\sum_h s_h V^h(q, \sum_{i\neq 0} \theta_i^h(1-\tau)\pi^i(p) + m) - \nu^T[\sum_h x^h(q, m + \sum_{i\neq 0} \theta_i^h(1-\tau)\pi^i(p)) - \sum_{i\neq 0} y^i(p) - y^0(p^0)]. \tag{7.2}$$

Assuming interior solutions and differentiability assumptions and exploiting the Roy's theorem, the first-order conditions are

$$-\sum_h s_h \nabla_{m^h} V^h x^{hT} - v^T \sum_h \nabla_q x^h = 0, \tag{7.3}$$

$$\sum_h s_h \nabla_{m^h} V^h \sum_{i\neq 0} \theta_i^h (1-\tau)\nabla_p^T \pi^i - v^T[\sum_h \nabla_{m^h} x^h \sum_{i\neq 0} \theta_i^h(1-\tau)\nabla_p^T\pi^i + \nabla_p \sum_{i\neq 0} y^i] = 0, \tag{7.4}$$

$$-\sum_h s_h \nabla_{m_h} V^h \sum_{i\neq 0} \theta_i^h \pi^i(p) + v^T \nabla_{m_h} x^h \sum_{i\neq 0} \theta_i^h \pi^i(p) = 0 \tag{7.5}$$

$$\sum_h s_h \nabla_{m^h} V^h - v^T \sum_h \nabla_{m^h} x^h = 0, \tag{7.6}$$

and

$$v^T \nabla_{p^0} y^0(p^0) = 0. \tag{7.7}$$

Recalling the differentiability and homogeneity property of $y^0(p^0)$ and the symmetry of the Jacobian of $y^0(p^0)$, (7.7) implies that $v^T = \delta p^{0T}$, for some $\delta > 0$. We can simplify (7.4) as

$$\sum_h [s_h \nabla_{m_h} V^h - \delta p^{0T} \nabla_{m^h} x^h] \sum_{i \neq 0} \theta_i^h (1-\tau) \nabla_p^T \pi^i + \delta p^{0T} \nabla_p \sum_{i \neq 0} y^i = 0 \tag{7.8}$$

(7.8) can be rewritten as

$$\alpha^T (1-\tau) \Theta \Delta(p) + \delta p^{0T} \nabla_p \sum_{i \neq 0} y^i = 0, \tag{7.9}$$

where

$$\alpha^T = \left[\, s_1 \nabla_{m_1} V^1 - \delta p^{0T} \nabla_{m^1} x^1 \quad \ldots \quad s_H \nabla_{m_H} V^H - p^{0T} \nabla_{m^H} x^H \,\right]. \tag{7.10}$$

(7.9) implies that production efficiency at the second-best is true if and only if $\alpha^T(1-\tau)\Theta\Delta(p)$ is zero. But, since $\Theta\Delta(p)$ is full row-ranked, this is true if and only if $\alpha^T = 0^{NT}$, that is,

$$\nabla_{m_h} V^h s_h = p^{0T} \nabla_{m^h} x^h, \ \forall h \tag{7.11}$$

or

$$\tau = 1. \tag{7.12}$$

(7.12) corresponds to the case of one-hundred percent profit taxation. (7.11), on the other hand corresponds to the case where the second-best is also the first-best. This is because, substituting from (7.11) into (7.3), we obtain

$$\delta p^{0T} \sum_h [\nabla_q x^h - \nabla_{m^h} x^h x^{hT}] = 0. \tag{7.13}$$

Which from the symmetry and homogeneity properties of the Slutsky matrix implies that p^{0T} is also proportional to q.

So, in this economy, we have production inefficiency at all second-best allocations either when the government can implement one-hundred percent profit taxation or when the second-best allocation is also a first-best.

8. Conclusion.

We highlighted some of the errors in the previous proofs about production efficiency in economies with firm-specific profit taxation. We then went on to suggest an alternative proof of this result. Production inefficiency implies that the private sector producer price vector and the shadow prices in the public sector (the latter reflect the true shadow prices in the economy) are not proportional. Lemma 2 proves the intuitive result that, in an institutional structure where private producers are price takers and maximize profits, the differences in the marginal rates of substitution in the private and public sectors can be exploited to construct small changes in the two price vectors that induce an increase in the aggregate output and aggregate income of the economy. The continuity and linear homogeneity of the profit function can be exploited (as in Step 2 of Theorems 1 and 2) to show that, if firm-specific profit taxation was allowed or if the matrices of private sector supplies and the ownership shares satisfies a certain rank condition, then there exist small changes in the producer prices that, in addition to increasing the aggregate output and income, also maintain the profit incomes of consumers and the aggregate profits in the economy at the levels of the status-quo. This means that the increase in aggregate income shows up as an increase in the tax and public sector incomes of the government, which can be used to change commodity taxes or to increase demogrant incomes of people in a Pareto improving way.

The mechanism suggests why this strategy does not work, generally, in most private ownership economies when there are restrictions on profit taxation. This is because, while a production inefficient status-quo suggests that there are changes in producer prices that can increase the aggregate output in the economy, all of the increased output may not, in general, become available to the government for designing Pareto improving changes in taxes and demogrant. Private ownership diverts some of the increased resources from the government coffers and puts it into the hands of consumers as profit incomes. But the private ownership structure could be such that it may lead to an inequitable distribution of profit incomes and a decrease in welfare of some consumers, which no government policy may be able to correct with the remaining resources, that is, there may exist no directions of change in the government policy instruments that are Pareto-improving, equilibrium preserving, and compatible with the existing private ownership structure.

The next set of questions are, in the class of private ownership economies, how generic is this phenomenon of production inefficiency, within a given economy what is the size and structure of second-best allocations that are production efficient, and if these sets are sizeable, then how do we recover shadow prices in these situations for cost-benefit tests, since producer prices no longer reflect the true shadow prices of the resources. These are further questions to be taken up in future research projects.

APPENDIX

Lemma 1: *Let Assumptions 1 and 2 hold and let $p \in B^c$ and $p^0 \in B^0$. Then $y^c(p) + y^0(p^0) \notin \hat{Y}$ implies that p is not proportional to p^0.*

Proof: Suppose there exist $\mu > 0$ such that $p^0 = \mu p$. The homogeneity property of supplies implies that $y^0(p^0) = y^0(p)$. Then the famous result by Koopmans on interchangeability of set summation and optimization implies that

$$\max_y \{p \cdot y | y \in Y^c + Y^0 = Y\} = \max_{y^c} \{p \cdot y^c | y^c \in Y^c\} + \max_{y^0} \{p \cdot y^0 | y^0 \in Y^0\} \tag{A.1}$$

and the solution to the left side of (A.1) $y(p)$ is exactly equal to $y^c(p)+y^0(p)$. But $y(p) \in \hat{Y}$. Therefore, $y^c(p) + y^0(p) \in \hat{Y}$, which is a contradiction. ■

Lemma 2: *Suppose Assumptions 1, 2, and 3 hold. Let $\bar{p} \in B^c$ and $\bar{p}^0 \in B^0$ such that $\bar{p}$ and $\bar{p}^0$ are not proportional. Let $\bar{y} := y^c(\bar{p}) + y^0(\bar{p}^0)$. Then there exist sequences $\{p^v\} \to \bar{p}$ and $\{p^{0^v}\} \to \bar{p}^0$ such that $y^c(p^v) + y^0(p^{0^v}) \gg \{\bar{y}\}$ for all big enough v.*

Proof: Smoothness of $\hat{Y}^c$ and $\hat{Y}^0$ implies that $H(\bar{p}, \bar{p} \cdot y^c(\bar{p}))$ and $H(\bar{p}^0, \bar{p}^0 \cdot y^0(\bar{p}^0))$ are unique supporting hyperplanes for Y^c and Y^0 at $y^c(\bar{p})$ and $y^0(\bar{p}^0)$, respectively.[19] Since $\bar{p}$ and $\bar{p}^0$ are not collinear, $H(\bar{p}, 0)$ is not a supporting hyperplane for $H_{\geq}(\bar{p}^0, 0)$ and $H(\bar{p}^0, 0)$ is not a supporting hyperplane for $H_{\geq}(\bar{p}, 0)$ at 0^N.

Step 1. We show that there exist $\Delta y^c \in \mathbf{R}^N$ and $\Delta y^0 \in \mathbf{R}^N$ such that the following is true: (the intuition becomes clear when one sees Figure 1 on the next page.)

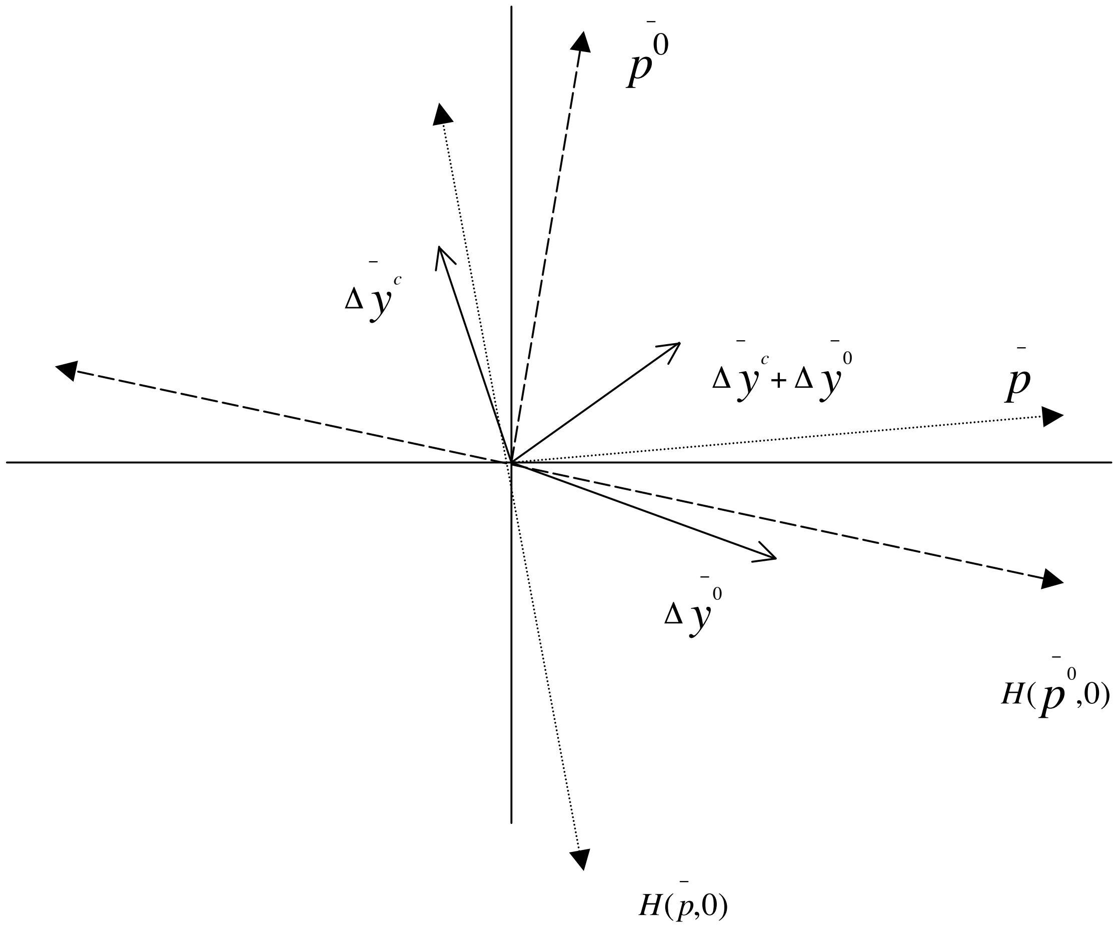
$$\begin{aligned} \Delta y^c \in H_{<}(\bar{p}, 0) \cap H_{\geq}(\bar{p}^0, 0), \\ \Delta y^0 \in H_{<}(\bar{p}^0, 0) \cap H_{\geq}(\bar{p}, 0), \text{ and} \\ \Delta y^c + \Delta y^0 > 0^N. \end{aligned} \tag{A.2}$$

Define $\rho = \lambda \bar{p} + (1 - \lambda)\bar{p}^0$, where $\lambda \in (0, 1)$. We will show that there exist Δy^c and Δy^0 as described in first two lines of (A.2) such that $\Delta y^c + \Delta y^0 = \rho > 0$.

Since $\bar{p}$ and ρ are not colinear, $H(p, 0)$ is not a supporting hyperplane of $H(\rho, 0)$, that is, $H(\rho, 0) \not\subset H_{\geq}(\bar{p}, 0)$. Hence, there exists $a \in H(\rho, 0)$ such that $a \cdot \bar{p} < 0$ and $a.\rho = 0$. Consider $\alpha a + (1 - \alpha)\rho$, where $\alpha \in (0, 1)$. Then

$$[\alpha a + (1 - \alpha)\rho] \cdot \rho = (1 - \alpha)\rho \cdot \rho > 0, \tag{A.3}$$

[19] We denote a hyperplane with normal p and constant a by $H(p, a)$ and its lower and strictly lower half-spaces by $H_{\leq}(p, a)$ and $H_{<}(p, a)$, respectively. Similarly we can define the upper and strictly upper half-spaces of $H(p, a)$.



$\bar{p}^0$
$\Delta \bar{y}^c$
$\Delta \bar{y}^c + \Delta \bar{y}^0$
$\bar{p}$
$\Delta \bar{y}^0$
$H(\bar{p}^0,0)$
$H(\bar{p},0)$

Figure 1

for all $\alpha \in (0,1)$. On the other hand, $[\alpha a + (1-\alpha)\rho] \cdot \bar{p} = \alpha \bar{p} \cdot a + (1-\alpha)\bar{p} \cdot \rho$. The first term is negative while the last term is non-negative. We can choose $\bar{\alpha}$ big enough so that

$$[\bar{\alpha} a + (1-\bar{\alpha})\rho] \cdot \bar{p} < 0. \tag{A.4}$$

$a \cdot \rho = 0$ implies that $a \cdot [\lambda \bar{p} + (1-\lambda)\bar{p}^0] = 0$ This implies that, since $a \cdot \bar{p} < 0$, we have $a \cdot \bar{p}^0 > 0$. So that $-a \cdot \bar{p}^0 < 0$. At the same time $-a \cdot \rho = 0$. Consider $\beta(-a) + (1-\beta)\rho$, where $\beta \in (0,1)$. Then

$$[\beta(-a) + (1-\beta)\rho] \cdot \rho = (1-\beta)\rho \cdot \rho > 0, \tag{A.5}$$

for all $\beta \in (0,1)$. On the other hand, $[\beta(-a) + (1-\beta)\rho] \cdot \bar{p}^0 = \beta \bar{p}^0 \cdot (-a) + (1-\beta)\bar{p}^0 \cdot \rho$. The first term is negative while the last term is non-negative. We can choose $\bar{\beta}$ big enough so that

$$[\bar{\beta}(-a) + (1-\bar{\beta})\rho] \cdot \bar{p}^0 < 0. \tag{A.6}$$

Define $\overset{*}{\alpha} = \max\{\bar{\alpha}, \bar{\beta}\}$. Define $\Delta y^c = \overset{*}{\alpha} a + (1-\overset{*}{\alpha})\rho$ and $\Delta y^0 = \overset{*}{\alpha}(-a) + (1-\overset{*}{\alpha})\rho$. Then $\Delta y^c + \Delta y^0 = 2(1-\overset{*}{\alpha})\rho > 0$.

Note that (A.4) and (A.6) imply that

$$\Delta y^c \cdot \bar{p} < 0 \tag{A.7}$$

and

$$\Delta y^0 \cdot \bar{p}^0 < 0. \tag{A.8}$$

(A.3) implies that $\Delta y^c \cdot \rho = \Delta y^c [\lambda \bar{p} + (1-\lambda)\bar{p}^0] = \lambda \Delta y^c \cdot \bar{p} + (1-\lambda)\Delta y^c \cdot \bar{p}^0 > 0$. Since the first term of the last inequality is negative from (A.7), the last term must be positive, that is, $\Delta y^c \cdot \bar{p}^0 > 0$. Hence, $\Delta y^c \in H_<(\bar{p}, 0) \cap H_\geq(\bar{p}^0, 0)$. Similarly, using (A.5) and (A.8) we can prove that $\Delta y^0 \in H_<(\bar{p}^0, 0) \cap H_\geq(\bar{p}, 0)$. Thus, Δy^c and Δy^0 satisfy all conditions of (A.2).

This implies that $y^c(\bar{p}) + y^0(\bar{p}^0) + \Delta y^c + \Delta y^0 > \bar{y}$. Denote $y^c(\bar{p})$ by $\bar{y}^c$ and $y^0(\bar{p}^0)$ by $\bar{y}^0$. Since $\bar{y}^c$ and $\bar{y}^0$ belong to $\hat{Y}^c$ and $\hat{Y}^0$, Assumption 3 implies that $f^c(\bar{y}^c) = 0$ and $f^0(\bar{y}^0) = 0$. Recall that $\nabla f^c(\bar{y}^c)$ is defined as the linear mapping such that for all $\{h^v\} \to 0^N$, we have

$$\lim_{h^v \to 0} \frac{f^c(\bar{y}^c + h^v) - [f^c(\bar{y}^c) + \nabla f^c(\bar{y}^c) h^v]}{|h^v|} \equiv \lim_{h^v \to 0^N} \frac{e(h^v, \bar{y}^c)}{|h^v|} = 0, \tag{A.9}$$

where $e(h^v, \bar{y}^c) = f^c(\bar{y}^c + h^v) - [f^c(\bar{y}^c) + \nabla f^c(\bar{y}^c) h^v]$.

Step 2. We show that there exists $\gamma^c > 0$ such that $\nabla f^c(\bar{y}^c) = \gamma^c \bar{p}$. Take any point $y \in Y^c$ such that $y \neq \bar{y}^c$. Then, from the convexity of Y^c, $f^c(\bar{y}^c + t(y - \bar{y}^c)) \leq 0$ for all $t \in (0,1)$. Using (A.9) and the fact that $f^c(\bar{y}^c) = 0$, we have

$$\frac{\nabla f^c(\bar{y}^c)(y - \bar{y}^c)}{|\, y - \bar{y}^c \,|} = \lim_{t \to 0} \frac{f^c(\bar{y}^c + t(y - \bar{y}^c)) - f^c(\bar{y}^c)}{|\, y - \bar{y}^c \,| \, t} = \lim_{t \to 0} \frac{f^c(\bar{y}^c + t(y - \bar{y}^c))}{|\, y - \bar{y}^c \,| \, t} \leq 0. \tag{A.10}$$

Since this is true for all $y \in Y^c$, $\nabla f^c(\bar{y}^c)$ is a normal to a supporting hyperplane of Y^c at $\bar{y}^c$. Since, $\hat{Y}^c$ is smooth and $H(\bar{p}, \bar{p} \cdot \bar{y}^c)$ is also a supporting hyperplane of Y^c at $\bar{y}^c$, there must exist $\gamma^c > 0$ such that $\nabla f^c(\bar{y}^c) = \gamma^c \bar{p}$. Similarly, we can prove that there exists $\gamma^0 > 0$ such that $\nabla f^c(\bar{y}^0) = \gamma^0 \bar{p}^0$.

(A.10) implies that $\Delta y^c \cdot \nabla f^c(\bar{p}) \leq 0$ and $\Delta y^0 \cdot \nabla f^0(\bar{p}^0) \leq 0$. Choose a sequence $\{t^v\}$ such that $t^v \Delta y^c \to 0$, and $t^v > 0$ for all v.

Step 3. We now show that there exists v' such that for all $v > v'$, we have $y^{cv} := \bar{y}^c + t^v \Delta y^c \in Y^c$. (A.9) implies that for any scalar t, we have

$$\frac{f^c(\bar{y}^c + t\Delta y^c) - f^c(\bar{y}^c)}{|\,\Delta y^c\,|\, t} = \frac{\nabla f^c(\bar{y}^c) t \Delta y^c}{t\,|\,\Delta y^c\,|} + \frac{e(t\Delta y^c, \bar{y}^c)}{t\,|\,\Delta y^c\,|}. \tag{A.11}$$

Then

$$\lim_{t^v \to 0} \frac{f^c(\bar{y}^c + t^v\Delta y^c) - f^c(\bar{y}^c)}{|\,\Delta y^c\,|\, t^v} = \frac{\nabla f^c(\bar{y}^c) \Delta y^c}{|\,\Delta y^c\,|} + \lim_{t^v \to 0} \frac{e(t^v\Delta y^c, \bar{y}^c)}{t^v\,|\,\Delta y^c\,|}. \tag{A.12}$$

Since $\frac{\nabla f^c(\bar{y}^c)\Delta y^c}{|\,\Delta y^c\,|} < 0$ and $\lim_{t^v\to 0} \frac{e(t^v\Delta y^c, \bar{y}^c)}{t^v|\,\Delta y^c\,|} = 0$ (this follows from (A.9)), we find that for large enough v, $\frac{e(t^v\Delta y^c, \bar{y}^c)}{t^v|\,\Delta y^c\,|}$ are dominated by $\frac{\nabla f^c(\bar{y}^c)\Delta y^c}{|\,\Delta y^c\,|}$. Noting that $f^c(\bar{y}^c) = 0$, this implies that there exists v' such that for all $v > v'$ we have

$$\frac{f^c(\bar{y}^c + t^v\Delta y^c)}{|\,\Delta y^c\,|\, t^v} = \frac{\nabla f^c(\bar{y}^c) \Delta y^c}{|\,\Delta y^c\,|} + \frac{e(t^v\Delta y^c, \bar{y}^c)}{t^v\,|\,\Delta y^c\,|} < 0. \tag{A.13}$$

Hence, for all $v > v'$, we have $f^c(\bar{y}^c + t^v\Delta y^c) < 0$, and hence $y^{cv} := \bar{y}^c + t^v \Delta y^c \in Y^c$ for all $v > v'$. Similarly, we can prove that there exists v'' such that for all $v > v''$, we have $y^{0^v} := \bar{y}^0 + t^v \Delta y^0 \in Y^0$.

Step 4. We now show that there exist sequences $\{p^v\}$ and $\{p^{0^v}\}$, and a positive integer $\hat{v}$ such that for all $v > \hat{v}$, we have $y^c(p^v) + y^0(p^{0^v}) > \bar{y}^c + \bar{y}^0$.

Define $\hat{v} = \max\{v', v''\}$. For every $v > \hat{v}$, $y^{cv} \in Y^c$. It can there be shown that there are continuous maps $\kappa^c(y^{cv}) := \max_\kappa \{\kappa \geq 0 \big| [y^{cv} + \kappa \mathbf{1}^N] \in Y^c\}$ and $\kappa^0(y^{0^v}) := \max_\kappa \{\kappa \geq 0 \big| [y^{0^v} + \kappa \mathbf{1}^N] \in Y^0\}$.[20] For all $v > \hat{v}$, it is clear that (i) $y^{cv} + y^{0^v} \gg \bar{y}^c + \bar{y}^0$ and so $(y^{cv} + \kappa^c(y^{cv})\mathbf{1}^N) + (y^{0^v} + \kappa^0(y^{0^v})\mathbf{1}^N) \gg \bar{y}^c + \bar{y}^0$, (ii) $(y^{cv} + \kappa^c(y^{cv})\mathbf{1}^N)$ and $(y^{0^v} + \kappa^0(y^{0^v})\mathbf{1}^N)$ belong to $\hat{Y}^c$ and $\hat{Y}^0$, respectively, and (iii) $\{y^{cv} + \kappa^c(y^{cv})\mathbf{1}^N\} \to \bar{y}^c$ and $\{y^{0^v} + \kappa^0(y^{0^v})\mathbf{1}^N\} \to \bar{y}^0$. Define $p^v = \frac{1}{\gamma^c}\nabla f^c(y^{cv} + \kappa^c(y^{cv})\mathbf{1}^N)$ and $p^{0^v} = \frac{1}{\gamma^0}\nabla f^0(y^{0^v} + \kappa^0(y^{0^v})\mathbf{1}^N)$. The smoothness of functions f^c and f^0 imply that $\{p^v\} \to \bar{p}$ and $\{p^{0^v}\} \to \bar{p}^0$. Clearly, $y^c(p^v) = y^{cv} + \kappa^c(y^{cv})\mathbf{1}^N$ and $y^0(p^{0^v}) = y^{0^v} + \kappa^0(y^{0^v})\mathbf{1}^N$, so that for all $v > \hat{v}$, we have $y^c(p^v) + y^0(p^{0^v}) \gg \bar{y}^c + \bar{y}^0$.

Hence, for all $v > \hat{v}$, the conclusions of the Lemma follow for sequences $\{p^v\}$ and $\{p^{0^v}\}$. ■

[20] See Appendix of a working paper version of this paper.

Lemma 3: *Suppose Assumptions 1 and 4 hold, $\Theta \in \Omega$ is such that $\theta_i^h > 0$ for all $i \neq 0$ and for all h, and the rank of the matrix $\Theta\Delta(p)$ is one. Pick any $h' \in \{1, \ldots, H\}$. Then there exist non-negative scalars μ^h for all h such that for all $p \in \mathbf{R}_{++}^N$, we have*

$$\sum_{i \neq 0} \theta_i^h y^{iT}(p) = \mu^h \sum_{i \neq 0} \theta_i^{h'} y^{iT}(p). \tag{A.14}$$

Proof: First note that for all h and for any $p \gg 0^N$, none of the rows of the matrix $\Theta\Delta(p)$ can be zeros: This follows from Remark 1 and the fact that $\theta_i^h > 0$ for all $i \neq 0$ and for all h, as then $\sum_{i \neq 0} \theta_i^h y^{iT}(p)p = \sum_{i \neq 0} \theta_i^h \pi^i(p) > 0$. Under the rank assumption, it follows that any row of $\Theta\Delta(p)$ provides a continuous (with respect to p) basis vector for the space spanned by the rows of $\Theta\Delta(p)$. Choose row h' as the basis. This implies that for all h, there exist continuous functions $\mu^h : \mathbf{R}_{++}^N \to \mathbf{R}_+$ such that for all h, we have

$$\sum_{i \neq 0} \theta_i^h y^{iT}(p) = \mu^h(p) \sum_{i \neq 0} \theta_i^{h'} y^{iT}(p). \tag{A.15}$$

Since supply vectors are homogeneous of degree zero in p, (A.15) implies that $\mu^h(p)$ is also homogeneous of degree zero in p. Exploiting the Hotelling's Lemma, we can rewrite (A.15) as

$$\sum_{i \neq 0} \theta_i^h \nabla_p^T \pi^i(p) = \mu^h(p) \sum_{i \neq 0} \theta_i^{h'} \nabla_p^T \pi^i(p), \ \forall\ h. \tag{A.16}$$

Post multiplying both sides of (A.15) with p we obtain

$$\sum_{i \neq 0} \theta_i^h \pi^i(p) = \mu^h(p) \sum_{i \neq 0} \theta_i^{h'} \pi^i(p). \tag{A.17}$$

Differentiating (A.17) we obtain

$$\sum_h \theta_{i \neq 0}^h \nabla_p^T \pi^i(p) = \mu^h(p)[\sum_{i \neq 0} \theta_i^{h'} \nabla_p^T \pi^i(p)] + \nabla_p^T \mu^h(p)[\sum_h \theta_{i \neq 0}^{h'} \pi^i(p)], \ \forall h. \tag{A.18}$$

A comparison of (A.16) and (A.18) imply that for all h

$$\nabla_p^T \mu^h(p)[\sum_h \theta_{i \neq 0}^{h'} \pi^i(p)] = 0^{NT}. \tag{A.19}$$

Under the maintained assumptions, $\sum_h \theta_{i \neq 0}^{h'} \pi^i(p) \neq 0$. Therefore (A.19) is true if and only if $\nabla_p \mu^h(p) = 0^N$ for all h, that is, if and only if $\mu^h(p)$ is a constant function for all h. Hence, (A.15) becomes (A.14). ■

REFERENCES

Dasgupta, P. and J. Stiglitz, "On Optimal Taxation and Public Production" *The Review of Economic Studies* **39**, 1972, 87-103.

Diamond, P. and J. Mirrlees, "Optimal Taxation and Public Production I: Production Efficiency" *The American Economic Review* **61**, 1971, 8-27.

Diamond, P. and J. Mirrlees, "Optimal Taxation and Public Production II: Tax Rules" *The American Economic Review* **61**, 1971, 261-278.

Drèze J. and N. Stern, *The Theory of Cost Benefit Analysis* in Handbook of Public Economics, editors: A. J. Auerbach and M. Feldstein, **II**, Chapter 14, Elsevier Science Publishers, North Holland, 1987.

Hahn F., "on Optimum Taxation" *Journal of Economic Theory* **6**, 1973, 96-106.

Guesnerie, R., *A Contribution to the Pure Theory of Taxation*, Cambridge, UK: Cambridge University Press, 1995.

Little I. M. D. and J. Mirrlees, *Project Appraisal and Planning for Developing Countries*, Heinemann Educational Books, London, 1974.

Mirrlees J., "On Producer Taxation" *Review of Economic Studies* **39**, 1972, 105-111.

Ramsey F., "A Contribution to the Theory of Taxation" *The Economic Journal* **37**, 1927, 47-61.

Reinhorn L., "Production Efficiency and Excess Supply" A Draft, 2005.

Sadka E., "A Note on Producer Taxation" *Review of Economic Studies* **44**, 1977, 385-387.

Weymark J., "On Pareto Improving Price Changes" *Journal of Economic Theory* **19**, 1978, 294-320.