Production efficiency and excess supply

Leslie J. Reinhorn
Economics Department
University of Durham
23-26 Old Elvet
Durham DH1 3HY
United Kingdom

phone +44 191 334 6365
fax +44 191 334 6341
reinhorn@hotmail.com

1 November 2011

I thank Peter Hammond, Gareth Myles, John Weymark, and a referee for helpful comments. This paper was completed during a visit to Union College in Schenectady, New York. I thank Union's economics department for their hospitality. This research was supported by a grant from the British Academy.

Abstract

This paper demonstrates that intermediate goods should not be taxed even in the presence of dividend payments to households, thus clarifying previous results. We also find that optimal government policy in a second best world may include stockpiles of output — private supply exceeds private demand, and the government purchases the surplus. This may provide a possible explanation for some agricultural policies.

JEL Classification: H21
Keywords: production efficiency; excess supply; optimal taxation; non-tight equilibrium; price supports

1 Introduction

Diamond and Mirrlees (1971) show that when the government sets tax rates optimally, equilibrium is characterized by production efficiency: transactions between firms should be free from distortionary taxation. The present paper asks if the production efficiency result continues to hold when firms distribute economic profits to households, a feature that was absent from Diamond and Mirrlees. Since dividends provide a direct link from firms to households, the government may wish to impose distortionary taxes on firms in order to manipulate profits and thereby affect households' incomes in a socially desirable way.

This use of distortionary taxation can be avoided if a profits tax may be imposed directly.[1] Indeed, Hahn (1973), Mirrlees (1972), and Sadka (1977) allow firm-specific taxation of economic profits. With this instrument, the government can control each firm's level of dividend payments — e.g., any increase in a firm's pre-tax profits can be neutralized with an increase in the profits tax (and conversely). As a consequence, we can separate the effects of producer prices from consumer prices. Based on the work of these authors, the conventional wisdom has been that the production efficiency theorem remains valid even in the presence of pure profits. However, the literature has gaps. There are some technical obstacles that make it quite difficult to provide a complete proof of the theorem. This raises concerns about the validity of the result, and it calls into question the conventional wisdom. The need for a correct proof seems clear, and this is provided in section 4. The proof introduces some novel features that allow equilibrium dividend payments to adjust continuously in response to changes in commodity tax rates. Thus, the clever insights of Hahn, Mirrlees, and Sadka are confirmed, and one of the most significant results in public economics is firmly established.[2]

Even with the production efficiency theorem intact, there still may be unexpected consequences from optimal taxation in the presence of dividends.[3] In the process of proving the production efficiency theorem, we find that optimality may include government stockpiles — e.g., agricultural surplus. That is, optimal tax policy may influence prices in such a way that private aggregate supply exceeds private aggregate demand. The government then purchases the surplus and places it in a stockpile, generating utility for no one. By comparison, in standard general equilibrium theory without government, if supply exceeds demand in a

[1] Besley and Persson (2009) observe that "when powers to tax are sufficient, it is always optimal ... to maximize national income and use the tax system to redistribute it" (page 1228).

[2] Murty (2010) has also addressed this issue, taking a different approach.

[3] "Unexpected" in the sense that second best optima may have properties that appear counter-intuitive to an observer who uses first best intuition (Lipsey and Lancaster 1956-1957).

market, the price must be zero. This is no longer true when government is present and taxes are distortionary. Conditions may arise in which optimal policy creates intentional waste through the hoarding or stockpiling of output even while production is carried out efficiently. In the absence of optimal lump sum transfers, this may be a second best method for getting income into the hands of some agents, particularly those who are favored by the social welfare function. Section 5 provides an example with optimal excess supply. One may ask why the government does not simply give away the surplus. The answer is that a giveaway would lower the price for the good in question, thus hurting some influential agents (farmers, in the case of agricultural stockpiles). Also, a giveaway may have unwanted general equilibrium repercussions via income effects.

The government's purchases of surplus may seem rather Keynesian in nature since they have no direct effect on the utility of any household. But recall that the optimal tax policy leads to production efficiency, with or without excess supply. Hence the purchases are not Keynesian in the traditional sense — they are not undertaken to correct an inefficiency. Instead they are motivated by distributional objectives.

1.1 Background

This section presents in general terms the gap in the literature's proofs of the production efficiency theorem for economies with profits. Appendix A provides the fine detail.

The proof of the production efficiency theorem uses the contrapositive: given any initial tax equilibrium that is productively *in*efficient, we can find a new tax equilibrium that is welfare-superior to the initial one. Thus production inefficiency cannot be optimal.

The problems in the literature can be illustrated, at least in general terms, with diagrams. The two dimensions of the page cannot tell the whole story, but the basic idea should follow. Consider an initial productively inefficient equilibrium A illustrated in figure 1. The curve in the first panel is firm 1's production efficiency frontier. This firm is producing at point A1 on its frontier, generating positive profits; similarly for firm 2 in the second panel. The economy as a whole is represented in the third panel. The aggregate production frontier is labeled "aggregate." The consumer's offer curve is also shown. By adjusting tax rates, and hence consumer prices, the government can move the consumer anywhere along the offer curve. In this particular equilibrium, the third panel shows aggregate production and consumption at point A which is the sum of A1 and A2. Production and consumption are required to coincide since previous proofs have not permitted excess supply. (Though see footnote 7.) While each firm individually is operating on its efficiency frontier, aggregate production is

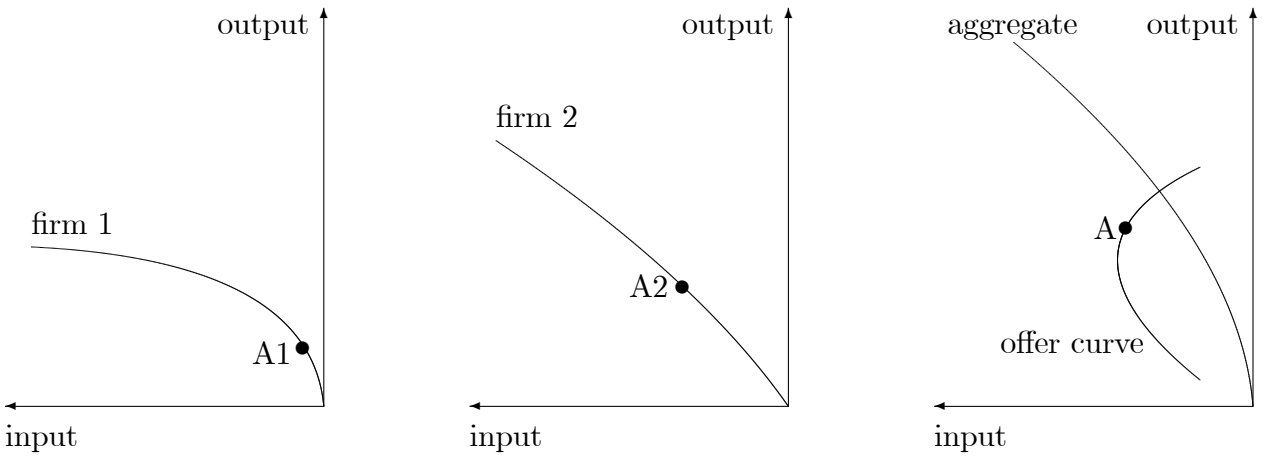

Figure 1: Initial equilibrium. Firm 1 produces at A1; firm 2 at A2. The sum of these points is at A in the third panel. Aggregate production is inefficient.

inefficient since firm 1's marginal product exceeds firm 2's.

The proof now identifies a new equilibrium B that is welfare-superior to A. This new equilibrium may be found by slightly reducing a tax rate from where it was in A. If consumers dislike this tax (and we should always be able to find a tax they do not like) welfare rises. The small tax change induces a small movement along the offer curve to point B in the third panel of figure 2. Since point A was in the interior of the aggregate production possibilities set, and since B is very close to A, B will be productively feasible — it will lie on or below the aggregate frontier. Since B is productively feasible it must be possible to divide up production between the two firms with the sum equal to B. Furthermore, since B is very close to A, the allocation of production across firms can be done so that each firm's production is very close to where it was in the initial equilibrium. Now here is where the argument runs into difficulty: since production for each firm has not moved very far, *neither have profits. And any slight change in profits can be offset with a slight change in the tax rates on profits to leave net dividends unaffected.* But this may not be the case. Instead, it may be that the only way to allocate the aggregate production point B across firms is as illustrated in the first two panels of figure 2. Firm 1's production point has moved only slightly from A1 to B1, and similarly for firm 2. Nonetheless, profits have moved discretely from positive to zero. The firms' new production points lie below their frontiers so they cannot be maximizers with positive profits.[4] This discrete change in profits causes a discrete change in the consumer's

[4]This simple approach with diagrams has its limitations. In figure 1 the slope at A1 differs from the slope at A2. By linear independence of the tangent lines, we should be able to find small movements along the

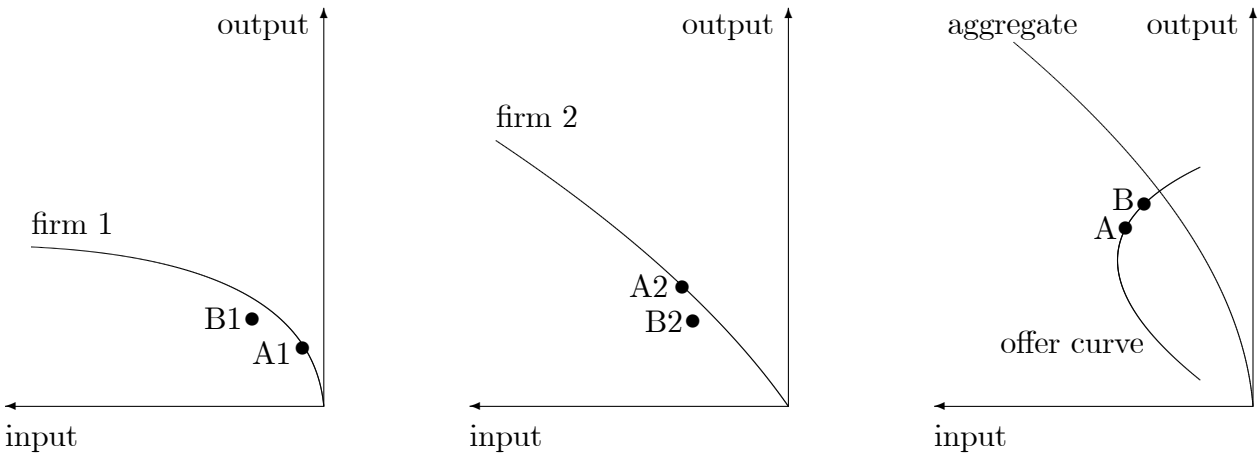

Figure 2: The proposed new equilibrium is labeled with Bs. The initial equilibrium is reproduced and is still labeled with As. New aggregate production and consumption are to be at B in the third panel. In order for production by the individual firms to sum to B, firm 1 produces at B1 in the first panel and firm 2 produces at B2 in the second panel. (The firm with the large marginal product expands and the firm with the small marginal product contracts.)

dividend income which cannot be offset with changes in the tax rates on profits. The result is then a shift in the offer curve (not illustrated) which is not accounted for by the proof in the literature. When the offer curve shifts, aggregate demand moves with it and demand no longer equals supply — the economy is no longer in equilibrium. There lies the problem.

The way I solve the problem is to allow for the possibility of excess supply. Then aggregate production can lie to the northeast of aggregate consumption. This is illustrated in figure 3. In the third panel, aggregate consumption stays at the same welfare-superior point B as in the previous figure 2. Production points for the "true" new equilibrium are labeled with Cs. In the first panel, firm 1 is now generating positive profits at C1 on its efficiency frontier to the northeast of B1. Since C1 is close to B1, and since B1 was close to A1, the profits at C1 are close to the profits at A1. A small change in firm 1's profits tax will leave net dividends exactly as they were in the initial equilibrium A; similarly for firm 2. Since the consumer's income from dividends has not changed, the offer curve remains in place, unlike the proposed construction with the Bs. In the third panel the aggregate production point is now at C, the sum of C1 and C2. This welfare-superior equilibrium as illustrated has excess supply.

firms' frontiers that add up to B. So we can maintain positive profits. However, in the more relevant case with many firms and many commodities this does not generalize.

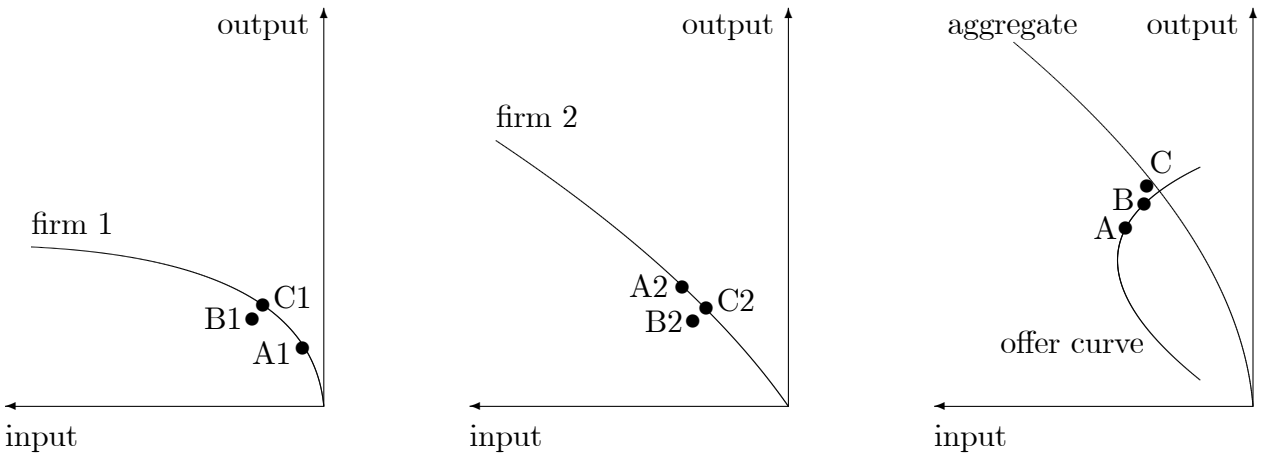

Figure 3: The true new equilibrium still has aggregate consumption at B in the third panel. Production moves to points labeled with Cs. In the first panel, firm 1 produces at C1 on its frontier to the northeast of B1. Similarly, firm 2 produces at C2. In the third panel, C is the sum of C1 and C2. It lies to the northeast of B, on or below the aggregate frontier.

Section 2 describes the model. Section 3 provides some preliminaries that are used to prove the main production efficiency result in section 4. Section 5 presents an example with optimal excess supply. The idea behind the example is as follows. One of the firms produces output in excess of consumer demand. The surplus could be eliminated if the firm simply produced less output from the same inputs. However, it is optimal for the firm to produce on its efficiency frontier since this generates positive profits which are distributed to households in a way that enhances social welfare. Section 6 contains concluding remarks.

2 Model

The model here is quite standard. After a brief description, notation and other details follow. Consumers are utility maximizing price takers. All consumers face the same prices. Taxes and subsidies are not modeled explicitly. Rather, they follow implicitly from the difference between consumer prices and producer prices.[5] Furthermore, different producers may face different prices. This allows for taxes and subsidies on intermediate goods — e.g., when the price paid by a retailer differs from the price received by a wholesaler, the difference is the

[5]It may be more appropriate to use buyer prices and seller prices rather than consumer prices and producer prices. However the use of the latter is completely standard in the literature. The two approaches are not equivalent. E.g., in a pure exchange economy there are no producer prices yet taxes may be imposed. One difficulty with the use of buyer prices and seller prices is the kink in consumers' budget sets.

tax or subsidy. It also allows for firm-specific tax rates on profits. Firms act in the interests of their shareholders, who can see through the corporate veil. Hence firms choose production levels to maximize after-tax profits. It follows that gross of tax prices have no bearing on firms' decisions, so in this paper any reference to producer prices will be net of all taxes. Production efficiency occurs when all firms face identical price ratios, or equivalently, when all firms face price vectors lying on the same line. This can be implemented by setting zero taxes on intermediate goods, while still permitting firm-specific tax rates on profits. Thus, when the production efficiency theorem applies, an optimizing government will choose not to tax intermediates even if it has the ability to do so.

Households are labeled $h = 1, \ldots, H$. Household h has consumption set $X_h \subset \mathbb{R}^n$ (net of endowment), utility function U_h, and lump sum income M_h. All households face the same vector of prices $\mathbf{q} \geq \mathbf{0}$.[6] Utility maximization results in net demand functions $\mathbf{x}_h(\mathbf{q}, M_h)$, defined on the domain where the maximum — which is assumed to be unique — exists. Aggregate net demand is $\mathbf{x}(\mathbf{q}, \mathbf{M}) := \sum_h \mathbf{x}_h(\mathbf{q}, M_h)$, defined on the domain where all of the $\mathbf{x}_h$s are defined.

Firms are labeled $f = 0, \ldots, F$. Firm f has convex net production set $Y_f \subset \mathbb{R}^n$. The aggregate production set is $Y := \sum_{f \geq 0} Y_f$. Firm 0 is the production unit for the public sector. Firms $f \geq 1$ are privately owned, profit maximizing, price takers. Producer prices are given by $\mathbf{p}_f$ and profits by π_f, both of which are net of producer taxes and taxes on profits. Assume $\mathbf{0} \in Y_f$ for $f \geq 1$, in which case $\pi_f \geq 0$. For each $f \geq 1$ define Y_f^+ to consist of all those production points that are capable of generating strictly positive profits. Specifically,

$$Y_f^+ := \{\mathbf{y}_f \in Y_f \,|\, \exists\ \mathbf{p} \in \mathbb{R}^n \text{ with } \mathbf{p} \cdot \mathbf{y}_f > 0 \ \&\ \mathbf{p} \cdot \mathbf{y}_f \geq \mathbf{p} \cdot \mathbf{y}\ \forall\ \mathbf{y} \in Y_f\}.$$

If we take a point in Y_f^+ and scale its supporting price vector up or down we can achieve any level of positive profits, as large or as small as we like. The process of scaling the price vector may be interpreted as an adjustment to the tax rate on profits. If we adjust too far we may get a rather impractical negative tax on profits, but this can always be avoided by re-normalizing the prices. The set Y_f^+ does not necessarily coincide with the boundary of Y_f. For instance, consider firms that have constant returns to scale.

The proportion of firm $f \geq 1$ owned by household h is $\theta_{hf} \geq 0$. Thus $\sum_h \theta_{hf} = 1$ for each $f \geq 1$. Let Θ be the $H \times F$ matrix with θ_{hf} in row h and column f. The government

[6] Notation for vector inequalities: $\mathbf{x} \geq \mathbf{y}$ if and only if all components of $\mathbf{x} - \mathbf{y}$ are non-negative; $\mathbf{x} \gg \mathbf{y}$ if and only if all components of $\mathbf{x} - \mathbf{y}$ are strictly positive.

imposes a head tax T (subsidy if negative). Therefore, $M_h = \sum_{f\geq 1} \theta_{hf}\pi_f - T$, or equivalently, $\mathbf{M} = \Theta\boldsymbol{\pi} - T\mathbf{1}$.

The government has a Bergson–Samuelson social welfare function W. Indirect social welfare is $V(\mathbf{q}, \mathbf{M}) := W\Big[\ldots, U_h\Big(\mathbf{x}_h(\mathbf{q}, M_h)\Big), \ldots\Big]$, which has the same domain as $\mathbf{x}$.

2.1 Definition. An *equilibrium* is a vector $(\mathbf{q}, \mathbf{M}, \mathbf{y}_0, \ldots, \mathbf{y}_F, \mathbf{p}_1, \ldots, \mathbf{p}_F, \boldsymbol{\pi}, T)$ that satisfies:

(a) $\mathbf{y}_f \in Y_f$ for each $f \geq 0$,

(b) $\pi_f = \mathbf{p}_f \cdot \mathbf{y}_f = \max\{\mathbf{p}_f \cdot \mathbf{y} \,|\, \mathbf{y} \in Y_f\}$ for each $f \geq 1$,

(c) $\mathbf{x}(\mathbf{q}, \mathbf{M}) \leq \sum_{f\geq 0} \mathbf{y}_f$,

(d) $\mathbf{M} = \Theta\boldsymbol{\pi} - T\mathbf{1}$.

Note the weak inequality in (c). This permits excess supply, which will be the focus of section 5. With regard to terminology, "excess supply" here is equivalent to "non-tight" equilibria in Guesnerie (1977). It also bears resemblance to the possibility of a government budget surplus in Berliant and Page (2001). In order to prove the results in sections 3 and 4 below, the weak inequality turns out to be crucial. The papers cited in appendix A do not permit excess supply and this leads to problems as outlined in section 1.1.[7]

If all households exhaust all their income then the government must satisfy its budget constraint with equality. This is just Walras' Law. In symbols, $\mathbf{q}\cdot\mathbf{x} + HT = \sum_{f\geq 1} \mathbf{p}_f \cdot \mathbf{y}_f$. An interpretation is that the government buys all output from private sector firms at producer prices then sells $\mathbf{x}$ to consumers at consumer prices, with added revenue effects from the head tax. Of course, this interpretation is excessively interventionist since the market can facilitate most transactions. However, the government does intervene directly to purchase the surplus, $\sum_{f\geq 0} \mathbf{y}_f - \mathbf{x}$. (Technically, any part of the surplus that the public sector produces using Y_0 is not "purchased." Rather, the inputs used to produce this output are purchased.) In this paper, excess supply refers to these residual purchases by the government.

[7] To be precise, Hahn (1973) on page 99 defines $Y_F(x)$ to permit excess supply. Yet the remainder of the paper does not seem to distinguish between aggregate net supply and aggregate net demand. Dixit (1987) on page 144 addresses the relationship between aggregate production inefficiency, excess supply, and free disposal. This relationship will be discussed further in section 5 below.

3 Almost production efficiency

If the production efficiency result holds, it can be stated in contrapositive form: For any equilibrium in which aggregate net output satisfies $\mathbf{y} \in \text{int}(Y)$, there exists another equilibrium with higher social welfare.[8] This section proves a weaker result (corollary 3.2). The condition $\mathbf{y} \in \text{int}(Y)$ is replaced with $\mathbf{y} \in \text{int}(\bar{Y})$ for a set $\bar{Y} \subset Y$ (not the closure of Y). This result will then be used in section 4 to prove the full production efficiency theorem.

Each equilibrium yields its own $\bar{Y}$. So consider an equilibrium, denoted by bars over variables. Then $\bar{Y}$ will consist of those aggregate production points that are capable of generating the same vector of profits as $\bar{\boldsymbol{\pi}}$. To construct $\bar{Y}$, first define $\bar{Y}_f$ for each $f \geq 1$ as follows. If $\bar{\pi}_f = 0$ then set $\bar{Y}_f := Y_f$. If $\bar{\pi}_f > 0$ then set $\bar{Y}_f := Y_f^+$. Thus all points in $\bar{Y}_f$ can preserve the sign of $\bar{\pi}_f$, and hence by scaling $\mathbf{p}_f$, can preserve the value of $\bar{\pi}_f$. Note that scaling $\mathbf{p}_f$ is equivalent to changing the profits tax rate for firm f. By construction, $\bar{\mathbf{y}}_f \in \bar{Y}_f$ for all $f \geq 1$. Now define $\bar{Y} := Y_0 + \sum_{f \geq 1} \bar{Y}_f$.

3.1 Theorem. *Assume* $\mathbf{x}$ *is a continuous function of* $\mathbf{q}$*, and* V *is a locally non-satiated function of* $\mathbf{q}$*. Consider an equilibrium denoted by bars over variables. For this equilibrium, define* $\bar{Y}$ *as above. If* $\mathbf{x}(\bar{\mathbf{q}}, \bar{\mathbf{M}}) \in \text{int}(\bar{Y} - \mathbb{R}^n_+)$ *then there exists another equilibrium — denoted by hats — with* $V(\hat{\mathbf{q}}, \hat{\mathbf{M}}) > V(\bar{\mathbf{q}}, \bar{\mathbf{M}})$.[9]

Proof. This is an application of familiar results (e.g., Mirrlees 1972). The hypotheses guarantee the existence of $\hat{\mathbf{q}}$ such that $V(\hat{\mathbf{q}}, \bar{\mathbf{M}}) > V(\bar{\mathbf{q}}, \bar{\mathbf{M}})$ and $\mathbf{x}(\hat{\mathbf{q}}, \bar{\mathbf{M}}) \in \bar{Y} - \mathbb{R}^n_+$, i.e., $\mathbf{x}(\hat{\mathbf{q}}, \bar{\mathbf{M}}) \leq \hat{\mathbf{y}}$ for some point $\hat{\mathbf{y}} \in \bar{Y}$. The new equilibrium will have $\hat{\mathbf{M}} = \bar{\mathbf{M}}$; hence, $\hat{\mathbf{x}} = \mathbf{x}(\hat{\mathbf{q}}, \bar{\mathbf{M}})$ and $\hat{V} = V(\hat{\mathbf{q}}, \bar{\mathbf{M}})$. Aggregate production will be at the point $\hat{\mathbf{y}} \in \bar{Y}$ just above. Also, the new head tax will be $\hat{T} = \bar{T}$. The proof will be complete if it is possible to allocate the aggregate production $\hat{\mathbf{y}} \in \bar{Y}$ across firms so that every private sector firm in the hat equilibrium generates the same after-tax profits as in the bar equilibrium. Then $\hat{M}_h$ equals income from profit shares minus the head tax, as required by part (d) of definition 2.1 (equilibrium). From the definition of $\bar{Y}$, it is indeed possible to allocate production in this way. (Though if $\bar{\pi}_f = 0$, it may be necessary to take $\hat{\mathbf{p}}_f = \mathbf{0}$: 100 percent taxation of profits.) ∎

3.2 Corollary. *Assume* $\mathbf{x}$ *is a continuous function of* $\mathbf{q}$*, and* V *is a locally non-satiated function of* $\mathbf{q}$*. Consider an equilibrium denoted by bars over variables. For this equilibrium,*

[8] That is, if production inefficiency is present then tax reform can lead to a welfare improvement. See Hammond and Sempere (1995) for a contribution to, and a review of, the tax/tariff reform literature.

[9] The Weymark condition (Diewert et al. 1989, Dixit 1987, Weymark 1979) is sufficient to guarantee that V is a locally non-satiated function of $\mathbf{q}$. That condition characterizes Pareto improving local changes in consumer prices.

define $\bar{Y}$ *as above. If* $\sum_{f\geq 0}\bar{\mathbf{y}}_f \in \operatorname{int}(\bar{Y})$ *then there exists another equilibrium — denoted by hats — with* $V(\hat{\mathbf{q}},\hat{\mathbf{M}}) > V(\bar{\mathbf{q}},\bar{\mathbf{M}})$.

Proof. Since $\sum_{f\geq 0}\bar{\mathbf{y}}_f \in \operatorname{int}(\bar{Y})$ it follows that $\mathbf{x}(\bar{\mathbf{q}},\bar{\mathbf{M}}) \in \operatorname{int}(\bar{Y}) - \mathbb{R}^n_+ \subset \operatorname{int}(\bar{Y} - \mathbb{R}^n_+)$. Now apply theorem 3.1. ∎

For an economy in which $\bar{Y} = Y$ corollary 3.2 yields full production efficiency. The Dasgupta–Stiglitz (1972) and Diamond–Mirrlees (1971) economies satisfy this condition.[10]

4 Smooth (enough) production frontiers

Corollary 4.5 below proves the claim made by Hahn (1973), Mirrlees (1972), and Sadka (1977) regarding the desirability of production efficiency. Specifically, if all private sector firms have smooth production frontiers, then any optimal tax equilibrium must be productively efficient. Assumption 4.1 formalizes the notion of a smooth (enough) production frontier.

4.1 Assumption. *If* $f \geq 1$ *and if* $\mathbf{y} \in Y_f^+$ *then there exists* $\epsilon > 0$ *such that* $Y_f \cap B_\epsilon(\mathbf{y}) \subset Y_f^+ - \mathbb{R}^n_+$ *where* $B_\epsilon(\mathbf{y})$ *is the open ball of radius* ϵ *centered at* $\mathbf{y}$.

Assumption 4.1 states that if a production point is sufficiently close to Y_f^+, there is a way to increase it (in the sense of $\mathbb{R}^n_+$) and enter Y_f^+. The first two panels of figure 3 previously illustrated the "increase" — each firm was able to move its production point northeast to its frontier where it was able to generate positive profits. Roughly, the assumption requires that if a firm's production frontier has any kinks, they must occur away from the outer edges of Y_f^+. Thus, the private sector production frontiers do not have to be perfectly smooth, only smooth enough.

The example of production *in*efficiency on page 107 of Mirrlees (1972) violates assumption 4.1. The essence of that example is illustrated in figure 4. Firm 1's constant returns to scale production frontier lies everywhere above firm 2's kinked production frontier. The kink violates assumption 4.1. Since firm 2's technology is dominated by firm 1's, it is productively inefficient for firm 2 to operate. However, firm 2 can generate profits while firm 1 cannot. Assume the economy has a household that needs dividends from these profits to survive. Then a utilitarian government will use taxes to keep firm 2 in operation with a price ratio

[10] The proof of theorem 3.1 made use of the possibility of excess supply — the possibility that condition (c) in definition 2.1 (equilibrium) holds with inequality. But even when excess supply is prohibited, as in much of the literature, corollary 3.2 remains true. With $\sum_{f\geq 0}\bar{\mathbf{y}}_f = \mathbf{x}(\bar{\mathbf{q}},\bar{\mathbf{M}})$, a minor modification to the proof of theorem 3.1 will prove corollary 3.2 directly.

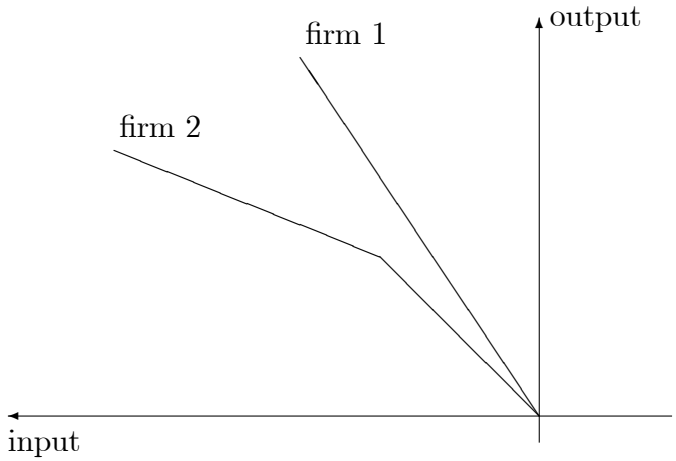

Figure 4: Production inefficiency. Firm 2 is less efficient than firm 1. But only firm 2 can generate positive profits.

that induces the firm, via profit maximization, to operate right at the kink point. Since the firm is inefficient we want it to be as small as possible, but with positive profits so the household survives. The kink does this for us — it establishes a smallest scale of operations for which profits are positive.

Mirrlees also considers an alternative case where firm 2 has a smooth and strictly concave production function (still dominated by firm 1), which would now satisfy assumption 4.1. In this case, if firm 2 produces any positive level of output, one could cut the scale of operations in half, say, and still generate positive profits. So no optimal tax equilibrium would exist: each equilibrium could be improved upon by price changes that cut inefficient firm 2's output in half and increase efficient firm 1's output correspondingly. The upshot is that the smoothness assumption guarantees any optimal tax equilibrium is productively efficient, but it does not guarantee the existence of an optimum.

If a solution to the optimal tax problem fails to exist, the government would then have to choose tax rates that are *almost* optimal. The equilibrium would not in general be productively efficient but we might want to know if it is *almost* productively efficient. Mirrlees states conditions which would apply to this case: "[A]ll producers either operate under constant returns, or obtain positive profit for any non-zero production under non-zero prices" (page 108). In this way, if a firm is kept in operation solely because its profits are socially desirable, the firm may be shrunk to an arbitrarily small size (hence, an arbitrarily small inefficiency) while still generating positive profits. These peculiarities arise out of situations

where dividend income is an indispensable part of redistribution. Since this is unlikely to be particularly important in practice, we shall move on.

Returning to assumption 4.1, we may find that it is difficult to verify in any given situation. However, in the more common case where firms' production sets are defined using continuously differentiable production functions, the assumption will be satisfied:

4.2 Theorem. *Let* $Y_f = \{(\mathbf{y}_f^o, \mathbf{y}_f^i) \in \mathbb{R}^n \,|\, G_f(\mathbf{y}_f^o, \mathbf{y}_f^i) \leq 0,\ \mathbf{y}_f^i \leq \mathbf{0}\}$. *The superscript o is for output, and i for input. Assume that* G_f *has a convex domain on which it is continuous, quasi-convex (convexity of* Y_f*), monotone non-decreasing (free disposal), and locally non-satiated. On* Y_f^+*, assume that* G_f *is continuously differentiable*[11] *with non-vanishing gradient. Then* Y_f *satisfies assumption 4.1.*

This theorem is proved in appendix B. Local non-satiation of G_f implies that the boundary of Y_f contains $\{(\mathbf{y}_f^o, \mathbf{y}_f^i) \in Y_f \,|\, G_f(\mathbf{y}_f^o, \mathbf{y}_f^i) = 0\}$. The partition between $\mathbf{y}^o$ and $\mathbf{y}^i$ is illustrated in the following example — the firm specific subscript f is omitted: $G(y^o, y_1^i, y_2^i) := y^o - (-y_1^i)^\alpha$ where $0 < \alpha \leq 1$. This firm uses y_1^i as an input to produce y^o, and it has no involvement in the market for the other input y_2^i. When $\alpha < 1$, Y^+ is the subset where $G = 0$ and $y^o > 0$.

The corollary to the following theorem will give the main production efficiency result.

4.3 Theorem. *Assume* $\mathbf{x}$ *is a continuous function of* $\mathbf{q}$*, and* V *is a locally non-satiated function of* $\mathbf{q}$*. Consider an equilibrium denoted by bars over variables. If assumption 4.1 is satisfied and if* $\mathbf{x}(\bar{\mathbf{q}}, \bar{\mathbf{M}}) \in \operatorname{int}(Y - \mathbb{R}_+^n)$ *then there exists another equilibrium — denoted by hats — with* $V(\hat{\mathbf{q}}, \hat{\mathbf{M}}) > V(\bar{\mathbf{q}}, \bar{\mathbf{M}})$.

The proof of this theorem draws on the following result (which does not require assumption 4.1). It extends corollary 3(a) of Hahn (1973).

4.4 Lemma. *Let* $\tilde{\mathbf{y}}_f \in Y_f$ *for* $f \geq 0$. *Set* $\tilde{\mathbf{y}} := \sum_{f \geq 0} \tilde{\mathbf{y}}_f$. *For any* $\boldsymbol{\epsilon} \gg \mathbf{0}$ *define* $K(\boldsymbol{\epsilon}) := Y_0 + \sum_{f \geq 1} \Big(Y_f \cap B_{\epsilon_f}(\tilde{\mathbf{y}}_f) \Big)$. *Note that* $\tilde{\mathbf{y}} \in K(\boldsymbol{\epsilon})$. *Let* $\tilde{\mathbf{x}} \leq \tilde{\mathbf{y}}$. *If* $\tilde{\mathbf{x}}$ *is a boundary point of* $K(\boldsymbol{\epsilon}) - \mathbb{R}_+^n$ *then* $\tilde{\mathbf{x}}$ *is also a boundary point of* $Y - \mathbb{R}_+^n$.

This lemma is proved in appendix C by adapting Hahn's argument. Observe that the notation $K(\boldsymbol{\epsilon})$ suppresses the dependence of this set on the particular production allocation. When this notation is used in the proof of theorem 4.3 below, it will refer to the production allocation in the bar equilibrium.

[11] That is, each point of Y_f^+ has a neighborhood on which G_f is continuously differentiable.

Proof of theorem 4.3. Apply the contrapositive of lemma 4.4 to the bar equilibrium in the statement of theorem 4.3. It follows that for any $\boldsymbol{\epsilon} \gg \mathbf{0}$, $\mathbf{x}(\bar{\mathbf{q}}, \bar{\mathbf{M}}) \in \operatorname{int}\Big(K(\boldsymbol{\epsilon}) - \mathbb{R}^n_+\Big)$. For present purposes, a particular choice of $\boldsymbol{\epsilon}$, denoted $\bar{\boldsymbol{\epsilon}}$, is required. To this end, for each $f \geq 1$ choose $\bar{\epsilon}_f > 0$ as follows. If $\bar{\pi}_f > 0$ then $\bar{\mathbf{y}}_f \in Y_f^+$. Hence choose $\bar{\epsilon}_f$ as provided for in assumption 4.1. If $\bar{\pi}_f = 0$ then set $\bar{\epsilon}_f := 1$.

Since $\mathbf{x}(\bar{\mathbf{q}}, \bar{\mathbf{M}}) \in \operatorname{int}\Big(K(\bar{\boldsymbol{\epsilon}}) - \mathbb{R}^n_+\Big)$, if $K(\bar{\boldsymbol{\epsilon}}) - \mathbb{R}^n_+ \subset \bar{Y} - \mathbb{R}^n_+$ then this theorem will follow from theorem 3.1. Hence we proceed to show $K(\bar{\boldsymbol{\epsilon}}) - \mathbb{R}^n_+ \subset \bar{Y} - \mathbb{R}^n_+$. Recall that the construction of $\bar{Y}$ distinguishes firms by their profits in the bar equilibrium so let $S_> := \{f \geq 1 \,|\, \bar{\pi}_f > 0\}$ and $S_= := \{f \geq 1 \,|\, \bar{\pi}_f = 0\}$. Then

$$\begin{aligned} K(\bar{\boldsymbol{\epsilon}}) - \mathbb{R}^n_+ &= \Big[Y_0 + \sum_{f \geq 1} \Big(Y_f \cap B_{\bar{\epsilon}_f}(\bar{\mathbf{y}}_f)\Big)\Big] - \mathbb{R}^n_+ \\ &\subset \Big[Y_0 + \sum_{S_>} \Big(Y_f^+ - \mathbb{R}^n_+\Big) + \sum_{S_=} Y_f\Big] - \mathbb{R}^n_+ \\ &= \Big[Y_0 + \sum_{S_>} Y_f^+ + \sum_{S_=} Y_f\Big] - \mathbb{R}^n_+ \end{aligned}$$

where the second line follows from assumption 4.1 and the choice of $\bar{\boldsymbol{\epsilon}}$. ∎

4.5 Corollary. *Assume* $\mathbf{x}$ *is a continuous function of* $\mathbf{q}$*, and* V *is a locally non-satiated function of* $\mathbf{q}$*. Consider an equilibrium denoted by bars over variables. If assumption 4.1 is satisfied and if* $\sum_{f \geq 0} \bar{\mathbf{y}}_f \in \operatorname{int}(Y)$ *then there exists another equilibrium — denoted by hats — with* $V(\hat{\mathbf{q}}, \hat{\mathbf{M}}) > V(\bar{\mathbf{q}}, \bar{\mathbf{M}})$.

Proof. See corollary 3.2.

Again, note that these proofs make use of the weak inequality in part (c) of definition 2.1 (equilibrium). Net demand by households can be less than net supply by firms, with the excess supply purchased by the government and stockpiled. For instance in theorem 4.3, aggregate production in the hat equilibrium will lie in the set $K(\bar{\boldsymbol{\epsilon}})$ while aggregate consumption will lie in $K(\bar{\boldsymbol{\epsilon}}) - \mathbb{R}^n_+$. If we change the definition and require demand to equal supply in all markets, it is not clear if the same type of proof could be used. But why require equality? We do actually observe government stockpiles of some commodities especially where price supports are in place. Furthermore, if we forbid stockpiles and impose equality we may cause a reduction in welfare. The next section provides a worked example in which a commodity is in excess supply at the optimal tax equilibrium.

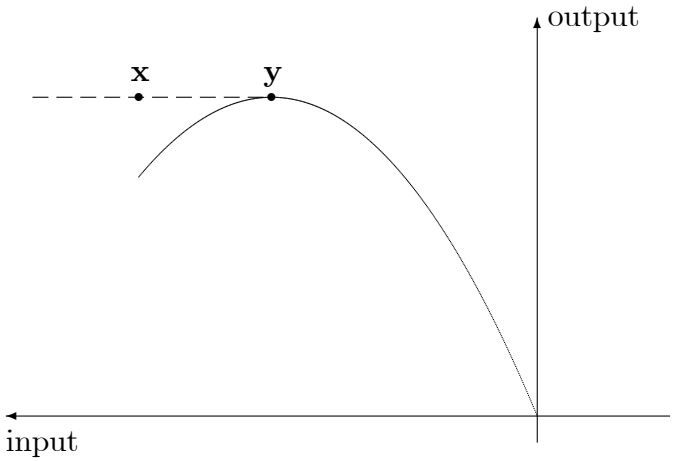

Figure 5: The production set is the region under the solid curve. The set $Y - \mathbb{R}^2_+$ also includes the region under the dashed line. There will be excess supply if net demand occurs at $\mathbf{x}$ and production at $\mathbf{y}$.

5 Excess supply

For any optimal tax equilibrium, theorem 4.3 above proved that aggregate net demand must lie on the boundary of $Y - \mathbb{R}^n_+$. This is essentially the tightest possible result since the actual prices, quantities, etc depend on the other data that describe the economy: the number of households, their preferences, their ownership shares, and the social welfare function.[12] In principle, any $\mathbf{x} \in \partial(Y - \mathbb{R}^n_+)$ can be supported as an optimum.

The possibility of optimal excess supply thus depends on the shape of Y. Specifically, it requires the existence of

$$\mathbf{x} \in \partial(Y - \mathbb{R}^n_+) \quad \text{and} \quad \mathbf{y} \in \partial Y \quad \text{with} \quad \mathbf{x} \leq \mathbf{y} \neq \mathbf{x}. \tag{1}$$

Figure 5 provides a crude illustration of this possibility.[13] The figure indicates that optimal excess supply requires flat segments in $\partial(Y - \mathbb{R}^n_+)$. This may be quite plausible when there are specialized factors of production (example 5.1 below). Flat segments may also appear when there is uncertainty, as in the technologies considered by Diamond (1967) where one input today yields several (state contingent) outputs tomorrow.

[12] The generic size of the set of second best tax equilibria is determined by the number of households (page 237 of Guesnerie 1995). Its position is determined by preferences (which here subsumes endowments) and by ownership shares. The social welfare function determines the selection from this set.

[13] Note, this is distinct from Guesnerie's (1977) *temporary* inefficiencies.

The following example takes a production technology that permits optimal excess supply and constructs the other ingredients to make this indeed optimal. The key feature of the example is that the commodity in excess supply does not satisfy the Diamond–Mirrlees condition. (Hereafter, DM.)[14] That is, some households are net suppliers of this commodity while other households are net demanders. Diamond and Mirrlees (1971) show that production efficiency is desirable if DM is satisfied (since DM implies local non-satiation of the indirect social welfare function). So if some, but not all, commodities have every household on the same side of the market then production efficiency and excess supply can both be desirable. Since the good in excess supply does not satisfy DM, lowering its consumer price in an effort to stimulate demand and reduce the surplus could have undesirable welfare consequences. Hence, it may be optimal to let the surplus be. This is what drives the example.

5.1 Example. There are four commodities: two types of completely specialized labor/leisure (ℓ and n), and two consumption goods (x and z). The economy is static. It would be easier to justify the complete specialization of labor in a dynamic model (e.g., I cannot supply labor services for time periods before I was born), but that would require a more elaborate structure. There are four households and two firms. There is no head tax. A head tax would give the government an extra degree of freedom that could be used to control households' incomes.[15] In order to limit the extent of this control it is simpler to eliminate the head tax rather than increase the number of households.

- Household 1 (type ℓ laborer) has utility function $U_1(\ell, x, z) = \log \ell + \log x + \log z$, which is written here as a function of consumption levels, though it could easily be converted into a function of net demand as in section 2. This household is endowed with 3/2 units of leisure. It has no ownership shares in either firm. The utility maximizing consumption levels satisfy $q_\ell \ell = q_x x = q_z z = q_\ell/2$ and the indirect utility function is $2 \log q_\ell - \log q_x - \log q_z + \text{constant}$.

- Household 2 (type n laborer) has utility function $U_2(n, x, z) = \log n + \log x + \log z$. It is endowed with 3/2 units of leisure and it has no ownership shares. The utility maximizing consumption levels satisfy $q_n n = q_x x = q_z z = q_n/2$ and the indirect utility function is $2 \log q_n - \log q_x - \log q_z + \text{constant}$.

[14] The DM condition, stated in theorem 4 on page 23 of Diamond and Mirrlees (1971), is the following: there exists a commodity for which every household is on the same side of the market. So DM is satisfied if all households are net suppliers of some commodity. It is also satisfied if all households are net demanders of some commodity.

[15] If the government has full control over all households' incomes, the outcome will be first best. And if preferences are strictly monotone, the first best cannot have excess supply.

- Household 3 (trader) has utility function $U_3(x,z) = \log x + \log z$, and it is endowed with 1 unit of good x. It has no ownership shares and it supplies no labor. The utility maximizing consumption levels satisfy $q_x x = q_z z = q_x/2$ and the indirect utility function is $\log q_x - \log q_z +$ constant. Note that household 3 receives the consumer price q_x for its net sales of good x, which could differ from the producer price p_x received by a firm. We can justify this if $p_x > q_x$ (a subsidy) since it is not practical to subsidize household to household transactions — it would bankrupt the government. The example does in fact allow for the possibility that $p_x > q_x$ at the optimum.

- Household 4 (capitalist) has utility function $U_4(x,z) = \log x + \log z$, and it has no endowment. It owns both firms, which yields total profits π. It supplies no labor to either firm. The utility maximizing consumption levels satisfy $q_x x = q_z z = \pi/2$ and the indirect utility function is $2\log\pi - \log q_x - \log q_z +$ constant.

- The government is not an active producer; $Y_0 = \{\mathbf{0}\}$.

- For firm 1, $Y_1 = \{(L,N,X,Z) \,|\, L \leq 0\,,\, N = 0\,,\, X \leq F(-L)\,,\, Z = 0\}$. This firm produces good x from type ℓ labor using a strictly increasing, strictly concave, smooth production function F.

- For firm 2, $Y_2 = \{(L,N,X,Z) \,|\, L = 0\,,\, N \leq 0\,,\, X = 0\,,\, Z \leq -N/2\}$. This firm produces good z from type n labor using a linear technology. It generates zero profits.

- The direct social welfare function is $W = U_1 + U_2 + 5U_3 + U_4$.

The government's problem is to maximize indirect social welfare subject to the weak inequalities for market clearing for each of the four commodities. If the level of production for firm 1 leads to excess supply, then the market clearing conditions for type ℓ labor and good x will not bind. Then the government's problem is to choose $\mathbf{q}$ and π to

$$\begin{aligned} \text{maximize} \quad & 2\log q_\ell + 2\log q_n + 2\log q_x - 8\log q_z + 2\log\pi \\ \text{subject to} \quad & q_\ell + q_n + q_x + \pi \leq q_z. \end{aligned}$$

The constraint incorporates the market clearing condition for good z, the production constraint for firm 2, and the market clearing condition for type n labor. This problem is homogeneous of degree zero in $(\mathbf{q},\pi)$, so normalize $q_z = 1$. Then the solution is $q_\ell = q_n = q_x = \pi = 1/4$.

At the optimal prices and profits, the supply of type ℓ labor is 1 and the aggregate net demand for good x is 1. To complete the example, choose the production function F for firm 1 so that $F(1) > 1$. Since this firm pays out positive profits at the optimum, it must produce on its efficiency frontier. Thus, there will be excess supply equal to $F(1) - 1$ units in the market for good x, which the government purchases. Alternatively, there could be excess supply in the market for type ℓ labor, which must be paid the wage $q_\ell = 1/4$ by the government.

We can say the following about optimal producer prices. Firm 2 with its linear technology must face the relative price $p_z/p_n = 2$, but p_z and p_n are not determined individually. If firm 1 uses one unit of type ℓ labor to produce $F(1) > 1$ units of good x at the optimum then its first order condition is $p_x F'(1) = p_\ell$ and its profit equation is $1/4 = \pi = p_x F(1) - p_\ell 1$. So $p_x = 0.25/[F(1) - F'(1)]$. We can choose the production function so that the producer price p_x exceeds the consumer price $q_x = 1/4$ in which case good x is subsidized as was mentioned above in the description of household 3, the trader. ∎

5.2 Remark. The trader plays an integral role in the example. The other three households prefer small values for q_x. In fact, as $q_x \downarrow 0$ their utilities and their consumption of x explode. Obviously this cannot be consistent with excess supply of x. The trader, on the other hand, prefers large values of q_x. This lack of unanimity allows a range of possible outcomes (depending on social welfare weights), including excess supply. This is the essence of the earlier discussion regarding the DM condition. ∎

Observe that the setup for the example satisfies the hypotheses for corollary 4.5. Thus, the example illustrates a relationship between production efficiency and excess supply. If the sufficient conditions for production efficiency are satisfied then the market clearing condition must bind for at least one market. However, it does not have to bind for every market.

It may be possible to eliminate excess supply entirely, without reducing social welfare. In particular, if the government has free disposal ($Y_0 - \mathbb{R}^n_+ \subset Y_0$), or if a private firm with constant returns has free disposal, then any excess supply can simply be thrown out.[16] But there is no real distinction between excess supply and government free disposal. Nor is there

[16] Weymark (1981) shows that the aggregate production set is equal to the sum of the boundaries of the firms' production sets: $\sum_f Y_f = \sum_f \partial Y_f$. Thus, it may seem that the presence or absence of a firm with free disposal is irrelevant. However, this result does not distinguish between Y_f^+ and ∂Y_f. There may be cases in which it is possible to re-allocate production so that all firms produce on their boundaries, but in the process one firm's production vector moves from Y_f^+ to $\partial Y_f \setminus Y_f^+$. This could affect profits and dividends, and hence affect net demand and social welfare.

any real distinction between private free disposal and public ownership (since price must be zero). Thus free disposal may effectively re-label, rather than eliminate, excess supply.

6 Conclusion

Production efficiency continues to be a topic of general interest to economists (e.g., Keen and Wildasin 2004). In this paper I extend the Diamond–Mirrlees (1971) production efficiency theorem to economies with pure profits. The result requires that small changes in demand be accommodated by small changes in supply without disrupting the level of dividends paid to households. Previous analyses have had difficulty formalizing this continuity assumption. The obstacles are addressed here by taking a new approach to define smoothness of the production frontier. Furthermore, the analysis here allows for the possibility of excess supply, or, in the terminology of Guesnerie (1977), allows for non-tight equilibria.

Example 5.1 illustrates that excess supply may indeed be optimal. The example is static and deterministic, but the model of section 2 is general enough to include commodities indexed by time and state of nature. These generalizations do not alter the key criterion: If the production set has the necessary shape as described in (1) then excess supply may be present at an optimal tax equilibrium.

Recall that the government absorbs the excess supply by purchasing it at market prices. As mentioned at the end of example 5.1, this can be achieved either by buying up inputs or outputs. Either way, the purchases are not consumed by any household. Rather, they are stockpiled by the government. Although this sounds particularly inefficient, it may be optimal given the constraints faced by the government. So it is natural to ask if we could achieve a better outcome by relaxing those constraints and giving the government more flexible policy instruments. The answer is yes if those instruments include unrestricted nonlinear taxation. The idea is to change the shape of the budget set so at least one household can afford more of the stockpiled commodity, while at the same time all other markets continue to clear. This eliminates the surplus but without thwarting social welfare objectives. Appendix D gives a formal statement and proof. Thus the excess supply may be avoided in principle. However, in practice unrestricted nonlinear taxation is not feasible due to the information requirements — the government needs to know the amount of each commodity purchased by each household.

If nonlinear taxation is not the answer perhaps we could consider firm specific lump sum transfers. After all, the key issue that led to excess supply in figure 3 was the desire to get

firms back on their efficiency frontiers so they could pay out the same level of dividends as in the initial equilibrium in figure 1. If we can achieve this directly with lump sum transfers, there is no need to introduce excess supply. Just give each firm a transfer that exactly restores the initial dividends. With this instrument the production efficiency theorem can be proved without the need for smooth production frontiers (assumption 4.1) and without the need for excess supply. It might appear that this new instrument is no more difficult to implement than the model's firm specific taxation of profits. In fact there is a real world policy that has been implemented and which has lump sum features: paying farmers not to farm. But this policy is a notorious magnet for abuse and corruption. The same can be expected of any subsidy that is unrelated to the level of production: everyone will try to get a piece of it. Unfortunately this leaves us again with the open question of whether there are any feasible policy instruments that can eliminate the surplus without reducing welfare.

Appendix A The literature

This appendix extends section 1.1 and is intended for those who have read the literature and want to see precisely where the difficulties arise. Four papers are of particular significance — Dasgupta and Stiglitz (1972), Hahn (1973), Mirrlees (1972), and Sadka (1977) — all of which use models similar to the one described in section 2 above.

Part (b) of the proof of Hahn's corollary 3 is not correct. The mapping that takes production points to supporting prices is not continuous as claimed — it is identically zero on the interior of the production set. Under stated assumptions, continuity could be achieved by restricting this mapping to the boundary of the production set.[17] But then the result from part (a) of the proof would not be applicable unless one were willing to assume convexity of the boundary of the production set — which essentially implies a linear technology. Sadka makes the same error.

Hahn's proposition 4 is somewhat misleading. In order to prove feasibility of his Pareto superior point he should show that if before-tax profits (π_f) equal zero then after-tax profits (n_f) equal zero. I.e., a firm cannot distribute profits that do not exist. However, the proof only seems to require the converse: $\pi_f > 0$ implies $n_f > 0$ (the stated restriction against 100% taxation of profits).

As pointed out by Sadka, Mirrlees's claim on the bottom of page 106 is in error. Mirrlees proceeds to consider a special case on the top of page 108. There are two types of firms: (i) those that are incapable of generating positive profits (firms with constant returns) and (ii) those for which Y_f^+ is dense in the boundary of Y_f. This is very restrictive since it excludes production sets like $Y_f = \{(y_1, y_2, y_3) \,|\, y_1 \leq \sqrt{-y_2}, y_2 \leq 0, y_3 \leq 0\}$ in which the firm is not involved in the market for good 3. Clearly this firm is not of type (i). Nor is it of type (ii) since Y_f^+ excludes all of the boundary points where $y_3 = 0$ and $y_1 < \sqrt{-y_2}$ (strict inequality). In practice, most firms participate in relatively few markets so it would be desirable to go beyond the special case considered by Mirrlees. Corollary 4.5 above does this.

Consider now the paper by Dasgupta and Stiglitz. There is one key assumption: profits are always strictly positive. Formally, each firm is characterized by a differentiable function that maps a normalized price vector $\mathbf{p}$ to a net supply vector. The assumption is that the inner product of these two vectors is strictly positive. This is similar to Mirrlees's special

[17] One should be careful when referring to the "boundary" and "interior" of the production set. Consider a firm for which there exists one commodity that is neither an input nor an output. Then the production set is contained in a hyperplane, which has no interior. In such a case it may be more appropriate to consider the relative topology induced by the hyperplane and use the terms "relative boundary" and "relative interior."

case. Despite the limitations from using calculus methods, the argument in Dasgupta and Stiglitz can be made rigorous. This follows from corollary 3.2 above.

Appendix B Proof of theorem 4.2

For ease of notation, omit the firm subscript f. Let ∇G denote the gradient of G. In order to prove the theorem, the following two lemmas will be helpful. The theorem's hypotheses also apply to these lemmas.

B.1 Lemma. *Let G be differentiable at $(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)$ with $G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) = 0$. If $G(\mathbf{y}^o, \mathbf{y}^i) \leq 0$ then $\nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \cdot (\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \geq \nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \cdot (\mathbf{y}^o, \mathbf{y}^i)$.*

Lemma B.1 shows that ∇G can serve as a supporting price vector. A proof is given on page 780 of Arrow and Enthoven (1961). Lemma B.2 below characterizes all supporting price vectors at points where $\nabla G \neq \mathbf{0}$.

B.2 Lemma. *Let G be differentiable at $(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)$ with $\nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \neq \mathbf{0}$. If $\mathbf{p} \cdot (\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \geq \mathbf{p} \cdot (\mathbf{y}^o, \mathbf{y}^i)$ for all $(\mathbf{y}^o, \mathbf{y}^i) \in Y$ then $\mathbf{p} = \bar{\alpha} \nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) + (\mathbf{0}, \bar{\mathbf{\Lambda}}^i)$ with $\bar{\alpha} \geq 0$, $\bar{\mathbf{\Lambda}}^i \geq \mathbf{0}$, and $\bar{\mathbf{\Lambda}}^i \cdot \bar{\mathbf{y}}^i = 0$.*

Proof. By hypothesis, $(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)$ is a solution to the following constrained optimization problem:

$$\max_{(\mathbf{y}^o, \mathbf{y}^i)} \ \mathbf{p} \cdot (\mathbf{y}^o, \mathbf{y}^i) \quad \text{subject to} \quad G(\mathbf{y}^o, \mathbf{y}^i) \leq 0 \quad \text{and} \quad \mathbf{y}^i \leq \mathbf{0}.$$

The Lagrangian for this problem is $\mathcal{L} = \mathbf{p} \cdot (\mathbf{y}^o, \mathbf{y}^i) - \alpha G(\mathbf{y}^o, \mathbf{y}^i) - \mathbf{\Lambda}^i \cdot \mathbf{y}^i$. Since monotonicity of G implies $\nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \geq \mathbf{0}$, a constraint qualification is satisfied at $(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)$. That is, the only Lagrange multipliers $(\alpha, \mathbf{\Lambda}^i) \geq \mathbf{0}$ that satisfy $\alpha \nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) + (\mathbf{0}, \mathbf{\Lambda}^i) = \mathbf{0}$ are $(\alpha, \mathbf{\Lambda}^i) = \mathbf{0}$. (Recall that $\nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \neq \mathbf{0}$.) Therefore, the Kuhn–Tucker conditions must be satisfied, and these conditions correspond to the conclusion of the lemma. ∎

To prove the theorem, let $(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \in Y^+$. The task is to find $\epsilon > 0$ that satisfies the condition in assumption 4.1. By definition of Y^+, lemma B.2 yields $\nabla G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \cdot (\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) > 0$. Since $\nabla G \geq \mathbf{0}$, this implies that for some output j, $\bar{y}^o_j > 0$ and $\partial G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)/\partial y^o_j > 0$. Without loss of generality, $j = 1$. Also, it follows that $G(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) = 0$. Otherwise profits could be raised by increasing y^o_1.

By the implicit function theorem, there exists a neighborhood $\mathcal{N}$ of $(\bar{\mathbf{y}}^o_{-1}, \bar{\mathbf{y}}^i)$ and a continuously differentiable function $g : \mathcal{N} \to \mathbb{R}$ such that $g(\bar{\mathbf{y}}^o_{-1}, \bar{\mathbf{y}}^i) = \bar{y}^o_1$ and $G\Big(g(\mathbf{y}^o_{-1}, \mathbf{y}^i), \mathbf{y}^o_{-1}, \mathbf{y}^i\Big)$

$\equiv 0$ on $\mathcal{N}$. Also, if $(\mathbf{y}^o_{-1}, \mathbf{y}^i) \in \mathcal{N}$ then $\partial G/\partial y^o_1$ remains strictly positive and ∇G remains continuous at $\Big(g(\mathbf{y}^o_{-1}, \mathbf{y}^i), \mathbf{y}^o_{-1}, \mathbf{y}^i\Big)$.

The mapping $(\mathbf{y}^o_{-1}, \mathbf{y}^i) \mapsto \nabla G(g, \mathbf{y}^o_{-1}, \mathbf{y}^i) \cdot (g, \mathbf{y}^o_{-1}, \mathbf{y}^i)$ is continuous on $\mathcal{N}$ where g is short for $g(\mathbf{y}^o_{-1}, \mathbf{y}^i)$. It maps $(\bar{\mathbf{y}}^o_{-1}, \bar{\mathbf{y}}^i)$ to a strictly positive number. Therefore, for sufficiently small $\epsilon > 0$ the open ball $B_\epsilon(\bar{\mathbf{y}}^o_{-1}, \bar{\mathbf{y}}^i)$ gets mapped into $\mathbb{R}_{++}$. The following lemma now confirms that this ϵ satisfies the condition in assumption 4.1.

B.3 Lemma. $Y \cap B_\epsilon(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i) \subset Y^+ - \mathbb{R}^n_+$.

Proof. Let $(\mathbf{y}^o, \mathbf{y}^i) \in Y \cap B_\epsilon(\bar{\mathbf{y}}^o, \bar{\mathbf{y}}^i)$. If $y^o_1 = g(\mathbf{y}^o_{-1}, \mathbf{y}^i)$ then by definition of ϵ, $\nabla G(g, \mathbf{y}^o_{-1}, \mathbf{y}^i) \cdot (g, \mathbf{y}^o_{-1}, \mathbf{y}^i) > 0$ so by lemma B.1 the proof is complete. Otherwise, the conditions $y^o_1 \neq g(\mathbf{y}^o_{-1}, \mathbf{y}^i)$ and $G(\mathbf{y}^o, \mathbf{y}^i) \leq 0$ imply $y^o_1 < g(\mathbf{y}^o_{-1}, \mathbf{y}^i)$ since G is monotone non-decreasing with $\partial G/\partial y^o_1 > 0$ at $\Big(g(\mathbf{y}^o_{-1}, \mathbf{y}^i), \mathbf{y}^o_{-1}, \mathbf{y}^i\Big)$. Thus $(\mathbf{y}^o, \mathbf{y}^i) = (g, \mathbf{y}^o_{-1}, \mathbf{y}^i) - (g - y^o_1, \mathbf{0}, \mathbf{0}) \in Y^+ - \mathbb{R}^n_+$. ∎

B.4 Remark. In some cases we may want to restrict the firm's ability to freely dispose goods that are neither inputs nor outputs. Then it may be more appropriate to consider production sets of the form $Y = \{(\mathbf{y}^o, \mathbf{y}^i, \mathbf{0}) \in \mathbb{R}^n \,|\, G(\mathbf{y}^o, \mathbf{y}^i) \leq 0,\ \mathbf{y}^i \leq \mathbf{0}\}$. Theorem 4.2 also holds for this Y.

Appendix C Proof of lemma 4.4

Since $\tilde{\mathbf{x}}$ is a boundary point of the convex set $K(\boldsymbol{\epsilon}) - \mathbb{R}^n_+$, there exists $\mathbf{a} \neq \mathbf{0}$ such that

$$\mathbf{a} \cdot \tilde{\mathbf{x}} \geq \mathbf{a} \cdot \mathbf{x} \quad \forall\ \mathbf{x} \in K(\boldsymbol{\epsilon}) - \mathbb{R}^n_+. \tag{2}$$

Since $\tilde{\mathbf{x}} \leq \tilde{\mathbf{y}}$ there exists $\tilde{\mathbf{b}} \geq \mathbf{0}$ such that $\tilde{\mathbf{x}} = \tilde{\mathbf{y}} - \tilde{\mathbf{b}}$.

Let $\mathbf{x}$ be a point in $Y - \mathbb{R}^n_+$ and suppose it has a representation $\mathbf{x} = \sum_{f \geq 0} \mathbf{y}_f - \mathbf{b}$ with $\mathbf{y}_f \in Y_f$ for each f and with $\mathbf{b} \geq \mathbf{0}$. For any $\lambda \in (0, 1)$ let $\mathbf{x}_\lambda := \lambda\tilde{\mathbf{x}} + (1-\lambda)\mathbf{x} = \sum_{f \geq 0}\Big(\lambda\tilde{\mathbf{y}}_f + (1-\lambda)\mathbf{y}_f\Big) - \Big(\lambda\tilde{\mathbf{b}} + (1-\lambda)\mathbf{b}\Big) \in Y - \mathbb{R}^n_+$. If λ is close to 1 then

$$\left\|\Big(\lambda\tilde{\mathbf{y}}_f + (1-\lambda)\mathbf{y}_f\Big) - \tilde{\mathbf{y}}_f\right\| = (1-\lambda)\|\mathbf{y}_f - \tilde{\mathbf{y}}_f\| < \epsilon_f \quad \forall\ f \geq 1.$$

Thus if λ is close to 1 then $\mathbf{x}_\lambda \in K(\boldsymbol{\epsilon}) - \mathbb{R}^n_+$, and hence from (2), $\mathbf{a} \cdot \tilde{\mathbf{x}} \geq \mathbf{a} \cdot \mathbf{x}_\lambda$. Since $1 - \lambda > 0$, the definition of $\mathbf{x}_\lambda$ and some simple algebra yields $\mathbf{a} \cdot \tilde{\mathbf{x}} \geq \mathbf{a} \cdot \mathbf{x}$. This is true for any $\mathbf{x} \in Y - \mathbb{R}^n_+$. Since $\mathbf{a} \neq \mathbf{0}$, it follows that $\tilde{\mathbf{x}}$ is a boundary point of $Y - \mathbb{R}^n_+$. ∎

Appendix D Nonlinear taxation

Nonlinear taxation was discussed briefly in the conclusion. To formalize this, replace the linear budget constraint $\mathbf{q} \cdot \mathbf{x}_h \leq \sum_f \theta_{hf}\pi_f - T$ with the more general constraint $Q(\mathbf{x}_h) \leq \sum_f \theta_{hf}\pi_f$ where the function Q is a policy choice for the government.

D.1 Theorem. *Consider an equilibrium in which there is excess supply of commodity j. Assume the following: (i) all households exhaust their budgets and one of the households has strictly greater income than all the others; (ii) the richest household's utility is strictly increasing in commodity j; (iii) social welfare is strictly increasing in the utility of the richest household. Then there exists another equilibrium with strictly greater social welfare.*

Proof. Let bars over variables denote the original equilibrium and let $\bar{z}_j$ be the amount of excess supply. Suppose household 1 has the strictly largest income in the bar equilibrium. Let hats over variables denote the welfare superior equilibrium. In this new equilibrium, production and profits remain as before. The new pricing function $\hat{Q}$ will coincide with $\bar{Q}$ except at one point: $\hat{Q}(\bar{\mathbf{x}}_1 + \bar{z}_j\mathbf{e}_j) = \bar{Q}(\bar{\mathbf{x}}_1)$ where $\mathbf{e}_j$ is the unit vector along the jth axis. By monotonicity of preferences, household 1 will now choose $\hat{\mathbf{x}}_1 = \bar{\mathbf{x}}_1 + \bar{z}_j\mathbf{e}_j$. All other households will leave their demand unchanged since the new price for $\bar{\mathbf{x}}_1 + \bar{z}_j\mathbf{e}_j$ is unaffordable. By definition of $\bar{z}_j$ these demands are feasible, and by the monotonicity assumptions social welfare has risen. ∎

References

[1] Arrow, Kenneth J. and Alain C. Enthoven (1961). "Quasi-concave programming." *Econometrica* **29**, 779–800.

[2] Berliant, Marcus and Frank H. Page, Jr. (2001). "Income taxes and the provision of public goods: Existence of an optimum." *Econometrica* **69**, 771–784.

[3] Besley, Timothy and Torsten Persson (2009). "The origins of state capacity: Property rights, taxation, and politics." *American Economic Review* **99**, 1218–1244.

[4] Dasgupta, Partha S. and Joseph E. Stiglitz (1972). "On optimal taxation and public production." *Review of Economic Studies* **39**, 87–103.

[5] Diamond, Peter A. (1967). "The role of a stock market in a general equilibrium model with technological uncertainty." *American Economic Review* **57**, 759–776.

[6] Diamond, Peter A. and James A. Mirrlees (1971). "Optimal taxation and public production I: production efficiency, and II: tax rules." *American Economic Review* **61**, 8–27 and 261–278.

[7] Diewert, W. Erwin, Arja H. Turunen-Red, and Alan D. Woodland (1989). "Productivity- and Pareto- improving changes in taxes and tariffs." *Review of Economic Studies* **56**, 199–215.

[8] Dixit, Avinash K. (1987). "On Pareto-improving redistributions of aggregate economic gains." *Journal of Economic Theory* **41**, 133–153.

[9] Guesnerie, Roger (1977). "On the direction of tax reform." *Journal of Public Economics* **7**, 179–202.

[10] Guesnerie, Roger (1995). *A Contribution to the Pure Theory of Taxation.* Cambridge: Cambridge University Press.

[11] Hahn, Frank H. (1973). "On optimum taxation." *Journal of Economic Theory* **6**, 96–106.

[12] Hammond, Peter J. and Jaime Sempere (1995). "Limits to the potential gains from economic integration and other supply side policies." *Economic Journal* **105**, 1180–1204.

[13] Keen, Michael and David Wildasin (2004). "Pareto-efficient international taxation." *American Economic Review* **94**, 259–275.

[14] Lipsey, Richard G. and Kelvin Lancaster (1956-1957). "The general theory of second best." *Review of Economic Studies* **24**, 11–32.

[15] Mirrlees, James A. (1972). "On producer taxation." *Review of Economic Studies* **39**, 105–111.

[16] Murty, Sushama (2010). "Constraints on profit income distribution and production efficiency in private ownership economies with Ramsey taxation." Working paper, University of Exeter.

[17] Sadka, Efraim (1977). "A note on producer taxation and public production." *Review of Economic Studies* **44**, 385–387.

[18] Weymark, John A. (1979). "A reconciliation of recent results in optimal taxation theory." *Journal of Public Economics* **12**, 171–189.

[19] Weymark, John A. (1981). "On sums of production set frontiers." *Review of Economic Studies* **48**, 179–183.

Notes

A construction that supports (1) as an optimum. There are 2 households and there is no head tax/subsidy. Household 1 fully owns all the firms. All social welfare weight is placed on household 2. Household 2's offer curve cuts $Y - \mathbb{R}^n_+$ at $\mathbf{x}$. I.e., the maximum of household 2's utility $U_2(\mathbf{x}_2)$ subject to $\mathbf{x}_2 \cdot \nabla U_2(\mathbf{x}_2) = 0$ and $\mathbf{x}_2 \in Y - \mathbb{R}^n_+$ occurs at $\mathbf{x}$. Household 1 has utility function $U_1(\mathbf{x}_1) := U_2(\mathbf{x}_1 + \mathbf{x})$, i.e., the same preferences over consumption as household 2 but a different endowment. Thus, when income is $M_1 = 0$ and consumer prices are proportional to $\nabla U_2(\mathbf{x})$, household 1 chooses autarky. The government taxes away all profits.

If we were willing to give up strict monotonicity, this construction would be trivial with Leontief preferences. I.e., let household 2's endowment be $\boldsymbol{\omega}_2$ with $\boldsymbol{\omega}_2 + \mathbf{x} \gg \mathbf{0}$ and with the Leontief kinks occurring where consumption has the proportions in $\boldsymbol{\omega}_2 + \mathbf{x}$. Let household 1 have the exact same preferences over consumption, with endowment at $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 + \mathbf{x}$. The outcome would be first best.

Former theorem 5.3 (which dealt with free disposal) may apply here. Since no profits are distributed, $\bar{Y} = Y$. And if Y has free disposal then there exists another optimum without excess supply. Nonetheless, the construction does demonstrate that (1) is sufficient to create the possibility of optimal excess supply. Furthermore, if the production point $\mathbf{y}$ does permit the distribution of positive profits, then other constructions may be possible for which former theorem 5.3 does not apply.

Buyer prices and seller prices. Consumer budget constraint $\mathbf{q} \cdot (\mathbf{x} - \boldsymbol{\omega})^+ \leq \mathbf{p} \cdot (\boldsymbol{\omega} - \mathbf{x})^+ - T$ where $\mathbf{x}$ is gross consumption not net demand.

Nonlinear taxation. In theorem D.1 consider what would happen if there were a tie for the largest income. Suppose the tie involves households 1 and 2. Consider $\hat{Q}$ that coincides with $\bar{Q}$ except as follows: $\hat{Q}(\bar{\mathbf{x}}_1 + \delta_1 \bar{z}_i \mathbf{e}_i) = \bar{Q}(\bar{\mathbf{x}}_1)$ and $\hat{Q}(\bar{\mathbf{x}}_2 + \delta_2 \bar{z}_i \mathbf{e}_i) = \bar{Q}(\bar{\mathbf{x}}_1)$ where the δs are non-negative with $0 < \delta_1 + \delta_2 \leq 1$. The goal might be to choose the δs to satisfy the following self-selection constraints:

$$U_1(\bar{\mathbf{x}}_1 + \delta_1 \bar{z}_i \mathbf{e}_i) \geq U_1(\bar{\mathbf{x}}_2 + \delta_2 \bar{z}_i \mathbf{e}_i) \qquad (\#)$$

and

$$U_2(\bar{\mathbf{x}}_2 + \delta_2 \bar{z}_i \mathbf{e}_i) \geq U_2(\bar{\mathbf{x}}_1 + \delta_1 \bar{z}_i \mathbf{e}_i). \qquad (*)$$

The problem is that this might not be possible. E.g., suppose both households were in a state of indifference in the bar equilibrium, $U_1(\bar{\mathbf{x}}_1) = U_1(\bar{\mathbf{x}}_2)$ and $U_2(\bar{\mathbf{x}}_2) = U_2(\bar{\mathbf{x}}_1)$. Suppose that household 1's marginal utility from commodity i is much greater at $\bar{\mathbf{x}}_2$ than at $\bar{\mathbf{x}}_1$. Then in order to satisfy $(\#)$, $\delta_1 > \delta_2$. Suppose also that household 2's marginal utility from commodity i is much greater at $\bar{\mathbf{x}}_1$ than at $\bar{\mathbf{x}}_2$. Then in order to satisfy $(*)$, $\delta_2 > \delta_1$. This appears problematic. However, the self-selection constraints will be satisfied if they are swapped. I.e., designate household 1 to demand $\bar{\mathbf{x}}_2 + \delta_2 \bar{z}_i \mathbf{e}_i$ and household 2 to demand $\bar{\mathbf{x}}_1 + \delta_1 \bar{z}_i \mathbf{e}_i$. In this way, the only change to aggregate demand is an increase in the demand for commodity i.

The difficulty arises when trying to generalize this argument for all possibilities. E.g., suppose both households are initially indifferent between $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ but that the marginal utilities have some other arrangement. Or suppose household 1 is indifferent between $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ while household 2 strictly prefers its own bundle to $\bar{\mathbf{x}}_1$. Etc. Then consider what happens if three or more households are tied for the largest income. It seems that the benefit from working through this problem just isn't worth the cost.