Optimal Inefficient Production*

Bas Jacobs**

August 18, 2015

Preliminary – Comments welcome

Abstract

This paper develops a model of optimal non-linear income and commodity taxation to analyze the desirability of aggregate production efficiency. In contrast to Diamond and Mirrlees (1971) individuals have individual-specific production technologies. It is demonstrated that the production efficiency theorem generally breaks down. Outputs of commodities should be taxed at higher (lower) rates if high- (low-)ability agents have a comparative advantage in producing them. In addition, outputs of commodities should be taxed relatively less when labor supply is more complementary to these outputs. Aggregate production efficiency is obtained only when the production technology for outputs does not depend on ability and is weakly separable from labor. The breakdown of the Diamond-Mirrlees production efficiency theorem has potentially important policy implications.

JEL code: H2
Key words: Diamond-Mirrlees production efficiency theorem, Atkinson-Stiglitz theorem, optimal non-linear income taxation, optimal commodity taxation

1 Introduction

Should the government distort production activities? That is the question raised by Diamond and Mirrlees (1971) in an article that is considered among the 20 most important papers of the *American Economic Review* during the last century. Diamond and Mirrlees demonstrated that it is optimal to operate an economy on the production possibilities frontier even in second-best situations where the government employs distortionary taxation.[1,2] This finding is often referred to as the *production efficiency theorem*. It provides the theoretical foundation for some very important policy prescriptions, such as the desirability of equal taxation of production sectors, the optimality of not taxing intermediate goods, the optimality free trade, the undesirability of

*The author thanks seminar participants of the CESifo Public Sector Area Conference 2014, and the UC Berkeley Public Economics Seminar, Björn Brügemann, Aart Gerritsen, Robin Boadway, Ray Rees, Emmanuel Saez, Danny Yagan, and especially Florian Scheuer and Josef Zweimüller for useful comments and suggestions. The usual disclaimer applies.

**Erasmus University Rotterdam, Tinbergen Institute, Netspar and CESifo. Address: Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Phone: +31–10–4081452/1441. Fax: +31–10–4089166. E-mail: bjacobs@ese.eur.nl. Homepage: http://people.few.eur.nl/bjacobs.

[1] Diamond and Mirrlees (1971) also note that two additional requirements are necessary: there need to be constant returns to scale in production or, if this is not the case, the government needs to have access to a perfect profit tax.

[2] Diamond and Mirrlees (1971) only considered linear taxes. Guesnerie and Seade (1982) demonstrated that the production efficiency theorem carries over in straightforward fashion to non-linear taxation.

source-based capital taxes like the corporate income tax, and it prescribes using market prices and discount rates in social cost-benefit analysis.[3]

This paper challenges the generality of the Diamond and Mirrlees (1971) production efficiency theorem. We develop a relatively standard optimal non-linear tax model with multiple commodities based on Diamond and Mirrlees (1971), Atkinson and Stiglitz (1976) and Mirrlees (1976). Individuals differ in their ability, which is private information. The model differs in one fundamental aspect from Diamond and Mirrlees (1971), Atkinson and Stiglitz (1976) and Mirrlees (1976): it is no longer assumed that every individual has access to the same (aggregate) production technology. Instead, it is assumed that every individual operates its own production technologies to produce different commodities. An individual's ability determines his/her productivity in transforming inputs in production into outputs for consumption. The main result of this paper is that when production technologies differ across individuals aggregate production efficiency is no longer socially desirable.

The fundamental reason that production is optimally inefficient is that the marginal rates of transformation between commodities provide valuable information on the hidden ability of individuals. If high-ability individuals have a comparative advantage in the production of a particular commodity, they will produce more of that commodity. The incentive-compatibility constraints associated with income redistribution can thus be relaxed by distorting production choices. Production of commodities should be taxed at higher (lower) rates when high-ability individuals have a stronger (weaker) comparative advantage in production of these commodities, i.e., when the marginal rates of transformation increase (decrease) with individual ability.

Moreover, we obtain a counterpart to Corlett and Hague (1953) and Atkinson and Stiglitz (1976) for the production side of the economy. Differential taxation of the production of commodities is shown to be optimal when labor supply has different elasticities of complementarity with the outputs of commodities. Intuitively, by distorting allocation of labor over various production activities, the tax burden can be shifted to production activities with a more inelastic labor demand, so that the total distortions in the labor market can be alleviated.[4]

The Diamond and Mirrlees (1971) production efficiency theorem assumes that every individual has access to identical technological possibilities to transform his/her inputs into outputs. In that case, the incentive-compatibility constraints do not depend on the production side of the economy and aggregate production efficiency prevails. Intuitively, no gain in redistribution can be obtained by distorting production if the marginal rates of transformation are the same for all agents, i.e., there is no comparative advantage in the labor market. Similarly, no reductions in labor market distortions can be obtained by distorting production, since labor demand elasticities are the same across sectors (i.e., infinite). Hence, production distortions should be avoided at all times. We derive the conditions under which aggregate production efficiency can be obtained. This depends on whether the production function for outputs is the same for everyone and is weakly separable from labor inputs. These conditions mirror the conditions

[3] Acemoglu, Tsyvinski, and Golosov (2008) show that the production efficiency theorem may even be valid in dynamic political-economy settings where there is no benevolent planner.

[4] According to the Atkinson and Stiglitz (1976) theorem there should optimally be no commodity-tax differentiation when individuals have identical and weakly separable utility functions. Our paper reveals that the Atkinson-Stiglitz theorem should be interpreted as a *consumption efficiency theorem*, which is the counterpart of the Diamond-Mirlees production efficiency theorem.

required for the Atkinson-Stiglitz theorem to have consumption efficiency.

Finally, Diamond and Mirrlees (1971) show that optimal tax expressions for both income and commodity taxation are the same in partial and general equilibrium and are independent from parameters from the production side. Saez (2004) refers to this finding as the 'Tax Formula' result. Our analysis demonstrates that this result is no longer applicable when individuals have different production technologies and production is optimally inefficient. Optimal tax formulae are then dependent on the parameters from the production side of the economy.

The policy implications of our analysis can be large. Our findings suggest that outputs from production activities in which high-ability agents have a comparative advantage should be taxed at higher rates compared to outputs from activities in which low-ability workers have a comparative advantage. Moreover, it is possible that minimum wages, industrial policies, trade restrictions and tariffs are socially desirable for that same reason. Furthermore, our model might explain why output of sectors with more (less) elastic labor demand should be taxed at relatively lower (higher) rates than other sectors. In addition, taxation of intermediate goods can be generally socially desirable. Finally, it is no longer guaranteed that using market prices in government production or social cost-benefit analysis is optimal. All in all, our analysis could provide a rationale as to why so many production inefficiencies are observed in the real world.

The rest of this paper is organized as follows. Section 2 discusses the relation to the existing literature. Section 3 develops the main argument Section 4 discusses the policy implications. Section 5 concludes.

2 Relation to the literature

Our analysis builds on several strands in the literature. First, numerous studies have explored the generality of the Diamond-Mirrlees production efficiency theorem. As already discussed in Diamond and Mirrlees (1971), the theorem is not applicable when not all transactions between firms and households can be taxed, including household production, see for example Stiglitz and Dasgupta (1971), Newberry (1986), and Kleven et al. (2000). Although these analyses are obviously important, we maintain the assumption from Diamond and Mirrlees (1971) that the government has access to a complete set of taxes on all outputs and commodities.

Second, Keen and Wildasin (2004) derive that the production efficiency does not generally apply in international settings, since international lump-sum transfers between countries are not available. These authors therefore cast doubts on the policy implications of the theorem for free trade, residence-based capital taxation and destination-based commodity taxation. We will assume a closed economy where such concerns do not arise.

Third, our findings are related to papers demonstrating that the production efficiency theorem is also not applicable when pure profits are not taxed at a 100 percent rate, see Diamond and Mirrlees (1971), Stiglitz and Dasgupta (1971) and Mirrlees (1972). In our model, the tax code is complete. The production efficiency theorem breaks down due to the presence of ability-specific quasi-rents.

Fourth, Naito (1999), Naito (2004) – analyzing commodity taxation – and Gaube (2005) – analyzing public-good provision – show that production efficiency breaks down when workers are imperfect substitutes in production. These authors use a Stiglitz (1982) 2-type model with two commodities being produced in two different sectors with constant-returns-to-scale production technologies. The Atkinson-Stiglitz theorem also no longer holds despite assuming weakly separable preferences. Their results are driven by general-equilibrium effects on prices and/or wages. Intuitively, changing factor prices affect the incentive-compatibility constraints, hence factor-price movements should be exploited for income redistribution. Production efficiency would be obtained in Naito (1999), Naito (2004) and Gaube (2005) when general-equilibrium effects on prices and wages are absent. Our findings differ importantly from these studies, since we assume constant prices. Moreover, we demonstrate that even when production is optimally distorted the Atkinson-Stiglitz theorem still applies, i.e., uniform commodity taxation is optimal, when preferences are identical and weakly separable.

Fifth, Saez (2004) follows up on Naito (1999) and develops an optimal tax model with occupational choice. Individuals can choose in which occupation/sector to work. Saez (2004) recovers both the Diamond-Mirrlees and Atkinson-Stiglitz theorems. In the long run, the relative supply to each occupation becomes infinitely elastic – even though labor types might be imperfect substitutes in production. This eliminates all differences in the production technology for individuals of different types. Since they only differ in their endowments of efficiency units of labor, all standard results are shown to be applicable again.

Sixth, Jacobs and Bovenberg (2011) refine the analysis Bovenberg and Jacobs (2005) to study optimal income taxation and optimal education subsidies in models with endogenous human capital formation and general specifications of the labor earnings functions. Human capital investment is the production of an intermediate good, which is then used in the final goods sector of the economy. Individuals with a higher ability are more productive in transforming human capital investment into labor earnings, that is high-ability individuals have a comparative advantage in skill formation. Jacobs and Bovenberg (2011) show that it is optimal to distort human capital formation, i.e., distort the demand for intermediate goods, for redistributional or efficiency reasons.

Seventh, Gomes, Lozachmeur, and Pavan (2014) analyze a 2-sector Roy-model of occupational choice with linear production technologies. These authors show that besides labor supply sectoral choice is distorted in second best, which they interpret as a violation of the Diamond-Mirrlees production efficiency theorem. However, one may question whether an occupational distortion, i.e. a distortion on labor supply on the extensive margin, should be interpreted as a production inefficiency given that all marginal rates of transformation are constant and equal to unity by definition. Gomes, Lozachmeur, and Pavan (2014) also demonstrate that the Atkinson-Stiglitz theorem breaks down and non-uniform commodity taxation is generally desirable.

3 Model

3.1 Individuals

There is a unit mass of individuals that are heterogeneous with respect to their one-dimensional productive ability $n \in \mathcal{N} \equiv [\underline{n}, \overline{n}]$, where $0 < \underline{n} < \overline{n} \leq \infty$. Ability is continuously distributed according to $H(n)$, which is the cumulative distribution function of n. $h(n)$ is the corresponding density function. Ability is private information. In contrast to Diamond and Mirrlees (1971) there is no aggregate production technology in which all workers are assumed to be perfect substitutes. Hence, ability is no longer equal to the number of efficiency units of labor. Ability n reflects an individual's productivity to transform inputs into outputs in production in a manner that will be made precise below.

All individuals have a strictly concave, continuous and twice differentiable utility function $u(\cdot)$. Individuals derive utility from I discrete commodities $x_i(n)$, indexed $i \in \mathcal{I} \equiv \{0, \cdots, I\}$. $x_0(n)$ is the untaxed, numéraire commodity. The vector of commodities consumed by a n-type individual is denoted by $\mathbf{x}(n) \equiv (x_0(n), x_1(n), \cdots, x_I(n))$. In addition, individuals derive disutility from supplying factor inputs $l_k(n)$, indexed $k \in \mathcal{K} \equiv \{0, \cdots, K\}$, where the vector of factor supplies is denoted by $\mathbf{l}(n) \equiv (l_0(n), l_1(n), \cdots, l_K(n))$. These factor inputs can be seen as different types of labor that are needed to produce the different commodities. For example each labor type k may refer to hours worked in performing a specific task or job. However, the factor supplies can be interpreted more broadly and may also include capital, land or other inputs in production. To allow for preference heterogeneity utility may depend on ability type n as in Mirrlees (1976). Hence, utility is written as:

$$u(n) \equiv u(\mathbf{x}(n), \mathbf{l}(n), n), \quad u_{x_i}, -u_{l_k} > 0, \quad u_{x_i x_i}, u_{l_k l_k} < 0, \quad \forall n. \tag{1}$$

All individuals can produce all commodities $\mathbf{x}(n)$, which are then publicly traded in markets. Prices of the x_i-goods are denoted by p_i. The vector of gross commodity prices is $\mathbf{p} \equiv (p_0, p_1, \cdots, p_I)$. The price of the numéraire good x_0 is normalized to unity without loss of generality, i.e. $p_0 = 1$. For simplicity, we confine the analysis to a partial-equilibrium setting where commodity prices $\mathbf{p}$ are fixed. The model can also be viewed as a small open economy where commodity prices $\mathbf{p}$ are determined in international goods markets.[5]

Each individual has access to a production technology to produce commodities. Outputs of commodities are designated by the vector $\mathbf{y}(n) \equiv (y_0(n), y_1(n), \cdots, y_I(n))$. All outputs $y_i(n)$ of the household are traded at gross market prices p_i. The production technology allows the individual to transform its K potential factor inputs into I potential outputs of each commodity. Individual-specific production sets are strictly convex and are denoted by $f(\cdot)$:

$$\begin{aligned} &f(\mathbf{y}(n), \mathbf{l}(n), n) = 0, \quad f_{y_i} > 0, \quad f_{l_k} < 0, \quad \forall n, i, k, \\ &\lim_{y_i \to 0} f_{y_i}(\cdot) = \infty, \quad \lim_{y_i \to \infty} f_{y_i}(\cdot) = 0, \quad \lim_{l_l \to 0} f_{l_k}(\cdot) = -\infty, \quad \lim_{l_l \to \infty} f_{l_k}(\cdot) = 0, \end{aligned} \tag{2}$$

The specification of the production set is general. If commodity i is an intermediate good it enters as a negative output y_i. These production technologies allow for *comparative advantage*: for a given amount of factor input individuals with a higher ability n can be relatively more

[5]In Naito (1999) and Naito (2004) commodity prices are endogenously determined in general equilibrium on goods markets of a closed economy. We abstract from this as general-equilibrium effects as such also determine the desirability of aggregate production efficiency. Our results demonstrate that production inefficiency is optimal even in the absence of general-equilibrium effects, in contrast to Naito (1999) and Naito (2004). As the analysis of optimal second-best production distortions in general equilibrium would become analytically quite complex, this extension is left for future research.

productive in the production of x_j-commodities than x_i-commodities, i.e., $f_{ny_j}/f_{y_j} < f_{ny_i}/f_{y_i}$ (and vice versa) for all j, i and n. The inequality means that the factor input requirement to produce one unit of output of x_j-commodities is smaller than the factor input requirement to produce one unit of output of x_i-commodities for individuals with a higher ability n. In addition, we assume *absolute advantage*: more able individuals, i.e. individuals with a higher n, are able to transform their factor input in more output of all commodities. Mathematically, this implies that $f_{ny_i} < 0$ for all n and i. Absolute advantage is required to ensure monotonicity, and hence implementability, of the optimal second-best allocation.

The assumption that everyone has a different production technology to produce each commodity is a relatively weak one. Diamond and Mirrlees (1971) assume that every individual has (access to) an *identical* production technology, which is a strong assumption. Diamond and Mirrlees (1971) is nested as a special case of our model, where the production functions are assumed to be identical for all individuals. In Mirrlees (1971) there is one output, one input, and a linear production technology. Hence, the production set is then given by: $f(y, l, n) \equiv y/n - l = 0$. Of course, this implies that gross earnings per worker y equal productivity times labor effort nl.

We assume equality in the production constraint. This implies that production takes place on the production possibilities frontier for each individual. We also impose Inada conditions on the production set for analytical tractability. These conditions ensure that each individual supplies labor effort in all jobs/tasks and produces all different commodities to some extent. This helps us avoid corner solutions and other, unimportant technical issues that arise if we would allow for extreme specialization patters. The analysis of corner solutions only distracts from the main messages of this paper. By suitable assumptions on the production set one can bring commodity demands or factor supplies arbitrarily close to zero.

Like in Mirrlees (1971), the government can neither verify individual factor supplies $l_k(n)$ nor their productive ability n. However, the government is able to verify output $y_i(n)$ in each production activity i and it can tax it accordingly. This is obviously a strong assumption, but it corresponds to Diamond and Mirrlees (1971) who assume that all production and consumption activities are verifiable by the government. Hence, all transactions between firms and households can be taxed. Indeed, when some intersectoral transactions cannot be verified, e.g. due to an informal sector, aggregate production efficiency ceases to be optimal, see also Stiglitz and Dasgupta (1971), Newberry (1986), and Kleven et al. (2000). By making these strong assumptions, we deliberately bias our findings towards the desirability of aggregate production efficiency.

The government levies non-linear taxes on the output volume of each commodity $T_i(p_i y_i(n))$, where the derivatives of the tax functions are assumed to be continuous, and denoted by $T_i'(p_i y_i(n)) \equiv \mathrm{d}T_i(p_i y_i(n))/\mathrm{d}(p_i y_i(n))$. In addition, the government is able to verify the consumption levels of all commodities. Hence, it can levy a set of non-linear ad valorem commodity taxes $t_i(p_i x_i(n))$ on all commodities, except for the numéraire commodity x_0 where $t_0 = 0$. Also, commodity tax functions are continuous and have derivatives $t_i'(p_i x_i(n)) \equiv \mathrm{d}t_i(p_i x_i(n))/\mathrm{d}(p_i x_i(n))$. It is not a priori clear that separate tax schedules on the outputs of each activity would support the optimal second-best allocation. However, under our assumptions the implementation through separate schedules on outputs and commodities works.[6] The individual budget constraint can thus be written as:

$$\sum_{i=0}^{I} p_i x_i(n) + t_i(p_i x_i(n)) = \sum_{i=0}^{I} p_i y_i(n) - T_i(p_i y_i(n)), \quad \forall n. \tag{3}$$

A number of things about the household budget constraint are noteworthy. First, in view of the non-linearity of the tax schedules $T_i(\cdot)$ and $t_i(\cdot)$ arbitrage seems profitable. However, since we assumed that all individual outputs and consumptions are verifiable to the government, such arbitrage is ruled out. Second, we explicitly allow for a different set of taxes on consumer and producer prices. Intuitively, the government would like to steer both the marginal rates of substitution in consumption and the marginal rates of transformation in production, since in our model each individual operates its own production technology. With a common, aggregate production technology, as in Diamond and Mirrlees (1971), the marginal rates of transformation do not vary by household type. We show later that this is the reason why aggregate production efficiency is optimal. Consequently, there would be no reason to have different non-linear schedules on different outputs, as these would only result in production distortions.

Households maximize utility (1) subject to their production set (2) and budget constraint (3). Setting up a Lagrangian for the individual's maximization problem, using $\lambda(n)$ and $\mu(n)$ as multipliers on the individual budget constraint and the production technology, gives:

$$\begin{aligned} \mathcal{L}(n) \equiv \quad & u\left(\mathbf{x}(n), \mathbf{l}(n), n\right) + \mu(n) f(\mathbf{y}(n), \mathbf{l}(n), n) \qquad (4)\\ & + \lambda(n)\left[\sum_{i=0}^{I} p_i y_i(n) - T_i(p_i y_i(n)) - \sum_{i=1}^{I} p_i x_i(n) + t_i(p_i x_i(n))\right], \quad \forall n. \end{aligned}$$

Necessary first-order conditions for an optimum are denoted by:

$$\frac{\partial \mathcal{L}(n)}{\partial x_i(n)} = u_{x_i}(\cdot) - \lambda(n) p_i (1 + t_i'(\cdot)) = 0, \quad \forall n, i, \tag{5}$$

$$\frac{\partial \mathcal{L}(n)}{\partial l_k(n)} = u_{l_k}(\cdot) + \mu(n) f_{l_k}(\cdot) = 0, \quad \forall n, k, \tag{6}$$

$$\frac{\partial \mathcal{L}(n)}{\partial y_i(n)} = \mu(n) f_{y_i}(\cdot) + \lambda(n) p_i (1 - T_i'(\cdot)) = 0, \quad \forall n, i. \tag{7}$$

These first-order conditions are not sufficient to describe the individual optimum despite the concavity assumptions on the utility and production functions. This is due to the non-linearity of the tax schedules. Below we will derive the second-order sufficiency conditions that need to be satisfied at the optimal second-best allocation.

These first-order conditions can be simplified to obtain the marginal rate of transformation $MRT_{ij}(n)$ of transforming x^i-goods into x^j-goods:

$$MRT_{ij}(n) \equiv \frac{(1 - T_j'(p_j y_j(n))) f_{y_i}(\mathbf{y}(n), \mathbf{l}(n), n)}{(1 - T_i'(p_i y_i(n))) f_{y_j}(\mathbf{y}(n), \mathbf{l}(n), n)} = \frac{p_i}{p_j}, \quad \forall n, i, j. \tag{8}$$

[6] Renes and Zoutman (2014) demonstrate that separate tax schedules indeed implement the optimal second-best allocation as long as there is a single-dimensional source of heterogeneity and there are no externalities. These conditions are satisfied in our model.

Hence, if production of outputs i and j are taxed at the same rates (i.e., $T_i'(\cdot) = T_j'(\cdot)$), then (individual) production efficiency is obtained among these activities, since the marginal rate of transformation equals the relative output price.

Similarly, we can find the marginal rate of substitution $MRS_{ij}(n)$ between x^i-goods and x^j-goods:

$$MRS_{ij}(n) \equiv \frac{u_{x_i}(\mathbf{x}(n), \mathbf{l}(n), n)}{u_{x_j}(\mathbf{x}(n), \mathbf{l}(n), n)} = \frac{(1 + t_i'(p_i x_i(n)))p_i}{(1 + t_j'(p_j x_j(n)))p_j}, \quad \forall n, i, j. \tag{9}$$

Hence, if marginal commodity taxes are equal ($t_i'(\cdot) = t_j'(\cdot)$), then individual consumption choices for these commodities are efficient.

Finally, the first-order conditions for factor supplies are denoted by:

$$\frac{u_{l_k}(\mathbf{x}(n), \mathbf{l}(n), n)}{u_{x_0}(\mathbf{x}(n), \mathbf{l}(n), n)} = (1 - T_0'(y_n^0)) \frac{f_{l_k}(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{y_0}(\mathbf{y}(n), \mathbf{l}(n), n)}, \quad \forall n, k. \tag{10}$$

Hence, factor supplies are not distorted if marginal tax rate on the output of the numéraire commodity is zero ($T_0'(\cdot) = 0$).

3.2 Social objectives and resource constraint

The government is assumed to maximize a utilitarian sum of utilities:

$$\int_{\mathcal{N}} u(\mathbf{x}(n), \mathbf{l}(n), n) \mathrm{d}H(n). \tag{11}$$

We assume that marginal utility of income is always declining with income. Diminishing private marginal utility of income yields a social preference for redistribution.[7] We assume that the government is purely redistributive, as there is no revenue requirement. The aggregate resource constraint of the economy is given by:

$$\int_{\mathcal{N}} \sum_{i=0}^{I} p_i y_i(n) - \sum_{i=0}^{I} p_i x_i(n) \mathrm{d}H(n) = 0. \tag{12}$$

The aggregate production constraint is given by:

$$\int_{\mathcal{N}} f(\mathbf{y}(n), \mathbf{l}(n), n) \mathrm{d}H(n) = 0. \tag{13}$$

The latter constraint implies that any production shortfall of one individual must be accompanied by a production surplus of another individual. Satisfaction of the resource and technology constraints and all the individual budget constraints implies that the government budget constraint will hold by Walras' law.

[7] We could adopt a generalized social welfare function or use Pareto weights, but doing so would not yield additional insights regarding the desirability of production (in)efficiency. Of course, the patterns of tax distortions are affected by the social objective.

3.3 First-best allocation

In order to interpret the second-best results derived below, we will first characterize the first-best allocation in which the government maximizes social welfare (11) subject to the aggregate resource constraint (12) and aggregate production constraint (13). By denoting by η as the multiplier associated with the aggregate resource constraint (12) and by κ the multiplier associated with the production constraint (13), the Lagrangian for this optimization problem can be written as:

$$\mathcal{L} \equiv \int_{\mathcal{N}} \left[u(\mathbf{x}(n), \mathbf{l}(n), n) + \eta \left(\sum_{i=0}^{I} p_i y_i(n) - \sum_{i=0}^{I} p_i x_i(n) \right) + \kappa f(\mathbf{y}(n), \mathbf{l}(n), n) \right] \mathrm{d}H(n), \quad (14)$$

Necessary and sufficient conditions for a first-best allocation are given by:

$$\frac{\partial \mathcal{L}}{\partial x_i(n)} = u_{x_i}(\cdot) - \eta p_i = 0, \quad \forall n, i, \quad (15)$$

$$\frac{\partial \mathcal{L}}{\partial y_i(n)} = \eta p_i + \kappa f_{y_i}(\cdot) = 0, \quad \forall n, i, \quad (16)$$

$$\frac{\partial \mathcal{L}}{\partial l_k(n)} = u_{l_k}(\cdot) + \kappa f_{l_k}(\cdot) = 0, \quad \forall n, k. \quad (17)$$

Hence, production efficiency is obtained since all marginal rates of transformation between all commodities x_i and x_j are equalized for all individual production decisions of all x_i commodities:

$$MRT_{ij}(n) \equiv \frac{f_{y_j}(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{y_i}(\mathbf{y}(n), \mathbf{l}(n), n)} = \frac{p_i}{p_j}, \quad \forall n, i, j. \quad (18)$$

Consumption efficiency is obtained since marginal rates of substitution between all commodities x_i and x_j are equalized for all individual consumption decisions of all x_i commodities:

$$MRS_{ij}(n) \equiv \frac{u_{x_i}(\mathbf{x}(n), \mathbf{l}(n), n)}{u_{x_j}(\mathbf{x}(n), \mathbf{l}(n), n)} = \frac{p_i}{p_j}, \quad \forall n, i, j. \quad (19)$$

Factor supplies are not distorted:

$$\frac{u_{l_k}(\mathbf{x}(n), \mathbf{l}(n), n)}{u_{x_0}(\mathbf{x}(n), \mathbf{l}(n), n)} = \frac{f_{l_k}(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{y_0}(\mathbf{y}(n), \mathbf{l}(n), n)}, \quad \forall n, k. \quad (20)$$

And, 'redistributional efficiency' is obtained, since all marginal utilities of consumption are equalized across all individuals:

$$u_{x_0}(\mathbf{x}(n), \mathbf{l}(n), n) = \eta, \quad \forall n. \quad (21)$$

3.4 Incentive compatibility

Given that ability n is private information, the first-best allocation cannot be obtained as it is not incentive compatible. Using a mechanism design approach, the optimal second-best allocation is characterized with the revelation principle. First, we derive the incentive-compatible direct mechanism. Second, we will decentralize this mechanism as an outcome of a competitive market using the non-linear tax schedules on outputs and commodities.

An allocation is said to be incentive compatible, when an individual of type n weakly prefers the bundle $\{\mathbf{x}(n), \mathbf{l}(n)\}$ of net consumption and factor supplies that the government intends for type n over the bundle $\{\mathbf{x}(n'), \mathbf{l}(n')\}$ intended for another type n'. Hence $u(n) = \max_{n'} u(\mathbf{x}(n), \mathbf{l}(n), n)$, $\forall n, n' \neq n \in \mathcal{N}$. As in Mirrlees (1971), we will apply the first-order approach to derive the second-best optimal allocation. The first-order incentive-compatibility constraint is derived in Lemma 1.

Lemma 1 *The first-order incentive-compatibility constraint is given by:*

$$\frac{\mathrm{d}u(n)}{\mathrm{d}n} = -u_{l_0}(\mathbf{x}(n), \mathbf{l}(n), n)\frac{f_n(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{l_0}(\mathbf{y}(n), \mathbf{l}(n), n)} + u_n(\mathbf{x}(n), \mathbf{l}(n), n), \quad \forall n. \tag{22}$$

Proof. Totally differentiating the utility function (1) gives – omitting the n-indices and function arguments:

$$\frac{\mathrm{d}u}{u_{x_0}} = \mathrm{d}x_0 + \sum_{i=1}^{I}\frac{u_{x_i}}{u_{x_0}}\mathrm{d}x_i + \sum_{k=1}^{K}\frac{u_{l_k}}{u_{x_0}}\mathrm{d}l_k + \frac{u_n}{u_{x_0}}\mathrm{d}n, \quad \forall n. \tag{23}$$

Totally differentiating the individual budget constraint (3) yields:

$$\mathrm{d}x_0 + \sum_{i=1}^{I} p_i(1+t_i')\mathrm{d}x_i = \sum_{i=0}^{I}(1-T_i')p_i\mathrm{d}y_i, \quad \forall n. \tag{24}$$

Next, use the first-order conditions for the individual problem (10) and (9) to substitute out the prices in (24):

$$\mathrm{d}x_0 + \sum_{i=1}^{I}\frac{u_{x_i}}{u_{x_0}}\mathrm{d}x_i = (1-T_0')\sum_{i=0}^{I}\frac{f_{y_i}}{f_{y_0}}\mathrm{d}y_i, \quad \forall n. \tag{25}$$

Totally differentiating the individual production set gives:

$$\sum_{i=1}^{I}\frac{f_{y_i}}{f_{y_0}}\mathrm{d}y_i + \sum_{k=0}^{K}\frac{f_{l_k}}{f_{y_0}}\mathrm{d}l_k + \frac{f_n}{f_{y_0}}\mathrm{d}n = 0, \quad \forall n. \tag{26}$$

Substitute (26) in (25) to eliminate $\sum_{i=1}^{I}\frac{f_{y_i}}{f_{y_0}}\mathrm{d}y_i$:

$$\mathrm{d}x_0 + \sum_{i=1}^{I}\frac{u_{x_i}}{u_{x_0}}\mathrm{d}x_i = -(1-T_0')\left(\sum_{k=0}^{K}\frac{f_{l_k}}{f_{y_0}}\mathrm{d}l_k + \frac{f_n}{f_{y_0}}\mathrm{d}n\right), \quad \forall n. \tag{27}$$

Substitute (27) in the totally differentiated utility function, use (10) and rewrite the resulting expression to establish Lemma 1. ■

In what follows, we will assume that the first-order approach is valid to characterize the optimal allocation. Second-order sufficiency conditions for utility maximization are respected under the constraint that is provided in Lemma 2.

Lemma 2 *We can invert the production sets $f(\mathbf{y}(n), \mathbf{l}(n), n) = 0$ to express the vector of inputs $\mathbf{l}(n)$ as a function of the vector of outputs $\mathbf{y}(n)$ and ability n: $\mathbf{l}(n) \equiv \phi(\mathbf{y}(n), n)$. Then, we* *rewrite the utility function (1) in terms of observables* $\mathbf{x}(n)$, $\mathbf{y}(n)$, *and ability* n *as*

$$u(n) \equiv u\left(\mathbf{x}(n), \mathbf{l}(n), n\right) = u\left(\mathbf{x}(n), \phi(n, \mathbf{y}(n)), n\right) \equiv V(x_0(n), \mathbf{X}(n), n), \tag{28}$$

where $\mathbf{X}(n) \equiv (\mathbf{x}_{-0}(n), \mathbf{y}(n))$ *and* $\mathbf{x}_{-0}(n) \equiv (x_1(n), x_2(n), \cdots, x_I(n))$ *is the vector of all commodities except the numéraire commodity. The following constraint on the Spence-Mirrlees and monotonicity conditions must hold at the optimal allocation:*

$$\frac{\mathrm{d}(V_{\mathbf{X}}/V_{x_0})}{\mathrm{d}n} \cdot \frac{\mathrm{d}\mathbf{X}'_n}{\mathrm{d}n} \geq 0. \tag{29}$$

Proof. See Mirrlees (1976, 334–335). ■

In what follows we will write consumption of the numéraire good $x_0(n)$ as a function of the allocation, that is: $x_0(n) \equiv x_0(\mathbf{x}_{-0}(n), \mathbf{l}(n), u(n), n)$, where $x_0(n)$ is obtained from inverting the utility function $u(n) = u(\mathbf{x}(n), \mathbf{l}(n), n)$. Derivatives of the consumption function $x_0(\cdot)$ are found using the implicit function theorem:

$$\frac{\partial x_0}{\partial l_k} = \frac{-u_{l_k}}{u_{x_0}} = -(1-T_0')\frac{f_{l_k}}{f_{y_0}}, \quad \frac{\partial x_0}{\partial x_i} = -\frac{u_{x_i}}{u_{x_0}} = -(1+t_i')p_i, \quad \frac{\partial x_0}{\partial u} = \frac{1}{u_{x_0}}, \quad \forall n. \tag{30}$$

By denoting η as the multiplier associated with the aggregate resource constraint (12), κ as the multiplier associated with the production constraint (13), $\theta(n)$ as the multiplier on the incentive-compatibility constraint (22), the Hamiltonian for maximizing social welfare can be formulated as:

$$\begin{aligned} \mathcal{H} \;\equiv\;& u(n)h(n) + \kappa f(\mathbf{y}(n), \mathbf{l}(n), n)h(n) \\ & + \eta\left(\sum_{i=0}^{I} p_i y_i(n) - x_0(\mathbf{x}_{-0}(n), \mathbf{l}(n), u(n), n) - \sum_{i=1}^{I} p_i x_i(n)\right) h(n) \\ & + \theta(n) u_{l_0}(x_0(\mathbf{x}_{-0}(n), \mathbf{l}(n), u(n), n), \mathbf{x}_{-0}(n), \mathbf{l}(n), n) \frac{f_n(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{l_0}(\mathbf{y}(n), \mathbf{l}(n), n)} \\ & - \theta(n) u_n(x_0(\mathbf{x}_{-0}(n), \mathbf{l}(n), u(n), n), \mathbf{x}_{-0}(n), \mathbf{l}(n), n), \quad \forall n. \end{aligned} \tag{31}$$

The necessary first-order and transversality conditions to characterize an optimal allocation are denoted by – omitting the indices n and the function arguments:

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial x_i} &= -\eta\left(p_i+\frac{\partial x_0}{\partial x_i}\right)h(n)+\theta\frac{f_n}{f_{l_0}}\left(u_{l_0x_i}+u_{l_0x_0}\frac{\partial x_0}{\partial x_i}\right) && (32)\\
&\quad -\theta\left(u_{nx_i}+u_{nx_0}\frac{\partial x_0}{\partial x_i}\right)=0,\ \forall n,i\\
\frac{\partial \mathcal{H}}{\partial y_i} &= \kappa f_{y_i}h(n)+\eta p_i h(n)+\theta u_{l_0}\left(\frac{f_{ny_i}f_{l_0}-f_{l_0y_i}f_n}{f_{l_0}^2}\right)=0,\quad \forall n,i, && (33)\\
\frac{\partial \mathcal{H}}{\partial l_k} &= \kappa f_{l_k}h(n)-\eta\frac{\partial x_0}{\partial l_k}h(n)+\theta\frac{f_n}{f_{l_0}}\left(u_{l_0l_k}+u_{l_0x_0}\frac{\partial x_0}{\partial l_k}\right) && (34)\\
&\quad +\theta u_{l_0}\left(\frac{f_{nl_k}f_{l_0}-f_{l_0l_k}f_n}{f_{l_0}^2}\right)-\theta u_n\left(u_{nl_k}+u_{nx_0}\frac{\partial x_0}{\partial l_k}\right)=0,\quad \forall n,i,\\
\frac{\partial \mathcal{H}}{\partial u} &= \left(1-\eta\frac{\partial x_0}{\partial u}\right)h(n)+\theta u_{l_0x_0}\frac{\partial x_0}{\partial u}\frac{f_n}{f_{l_0}}-\theta u_{nx_0}\frac{\partial x_0}{\partial u}=\frac{\mathrm{d}\theta}{\mathrm{d}n},\quad \forall n\neq \underline{n},\overline{n}, && (35)\\
&\lim_{n\to\underline{n}}\theta_n=\lim_{n\to\overline{n}}\theta_n=0. && (36)
\end{aligned}$$

3.5 Optimal consumption inefficiency

It will be useful to first derive the optimal consumption taxes, since the expressions for the optimal production taxes will yield a mirror image. The next Proposition demonstrates under which conditions it is optimal to distort consumption patters. Basically, this proposition generalizes Atkinson and Stiglitz (1976) and Mirrlees (1976) to a setting with production distortions.

Proposition 1 *The optimal non-linear marginal tax rates on demand for x_i-commodities are given by:*

$$\frac{t_i'(p_ix_i(n))}{1+t_i'(p_ix_i(n))}=\frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)}\left(\frac{\partial\ln(u_{x_i}(\cdot)/u_{x_0}(\cdot))}{\partial\ln n}-\frac{nf_n(\cdot)}{l_0f_{l_0}(\cdot)}\frac{\partial\ln(u_{x_i}(\cdot)/u_{x_0}(\cdot))}{\partial\ln l_0}\right),\quad\forall n,i. \tag{37}$$

Proof. Substitute the derivatives of x_0 in (30) into (32), rewrite using the first-order conditions (9) – omitting the indices and function arguments:

$$\frac{p_i-\frac{u_{x_i}}{u_{x_0}}}{\frac{u_{x_i}}{u_{x_0}}}=\frac{u_{x_0}\theta/\eta}{nh(n)}\frac{nf_n}{l_0f_{l_0}}\left(\frac{l_0u_{l_0x_i}}{u_{x_i}}-\frac{l_0u_{l_0x_0}}{u_{x_0}}\right)-\frac{u_{x_0}\theta/\eta}{nh(n)}\left(\frac{nu_{nx_i}}{u_{x_i}}-\frac{nu_{nx_0}}{u_{x_0}}\right). \tag{38}$$

Note that $\frac{l_0u_{l_0x_i}}{u_{x_i}}-\frac{l_0u_{l_0x_0}}{u_{x_0}}=\frac{\partial\ln(u_{x_i}/u_{x_0})}{\partial\ln l_0}$ and $\frac{nu_{nx_i}}{u_{x_i}}-\frac{nu_{nx_0}}{u_{x_0}}=\frac{\partial\ln(u_{x_i}/u_{x_0})}{\partial\ln n}$. Finally, the tax implementation uses the household's first-order condition from (9), i.e. $\frac{u_{x_i}}{u_{x_0}}=(1+t_i')p_i$, which can be used to eliminate $\frac{u_{x_i}}{u_{x_0}}$ to establish the proposition. ■

The left-hand side of equation (37) gives the non-linear marginal tax wedge on commodity x_i. The right-hand side of equation (37) gives the marginal benefits of taxing commodity x_i. The marginal benefits of (differential) commodity taxation are twofold. First, commodity taxation helps to complement the income tax to achieve the distributional goals of the government, as captured by the first term in brackets on the right-hand side. In particular, if high-ability individuals have a stronger (weaker) taste for commodity x_i than low-ability individuals do, then the marginal rate of substitution u_{x_i}/u_{x_0} between both commodities increases (declines) with n, so that $\frac{\partial \ln(u_{x_i}/u_{x_0})}{\partial \ln n} > 0$ (< 0). Consequently, commodity x_i needs to be taxed (subsidized). Of course, this distributional benefit also comes at a cost of distorting commodity demands. In the optimum, the government equates marginal redistributional benefits of commodity taxes and marginal distortions in commodity demands. This motive for differential commodity taxes is known since Mirrlees (1976), and has later been explored further by Saez (2002).

Second, differential commodity taxation is employed to alleviate the distortions of income taxation on labor supply, as the second term in brackets on the right-hand side demonstrates. Differential commodity taxation is optimal when the marginal rate of substitution between commodity x_i and the numéraire commodity x_0 varies with labor supply: $\frac{\partial \ln(u_{x_i}/u_{x_0})}{\partial \ln l_0} \neq 0$. That is, when good x_i is more (less) complementary to labor than good x_0 is, it will be optimal to subsidize (tax) x_i. Intuitively, by introducing a distortion in commodity demands, the government is able to alleviate some of the distortions in labor supply created by the non-linear income tax. See also Atkinson and Stiglitz (1976) and Jacobs and Boadway (2014).

The famous Atkinson-Stiglitz theorem is recovered when if the utility function is identical for all n and preferences are weakly separable between commodities and labor, so that $\frac{\partial \ln(u_{x_i}/u_{x_0})}{\partial \ln n} = \frac{\partial \ln(u_{x_i}/u_{x_0})}{\partial \ln l_0} = 0$. In that case, $t'(p_i x_i(n)) = 0$, since commodity taxes have no redistributional benefit over and above income taxes and commodity taxes are impotent to boost downward distorted labor supply. Hence, all commodities should be uniformly taxed (at zero rates). One may therefore interpret the Atkinson-Stiglitz theorem as the 'consumption efficiency theorem' of public finance.

Naito (1999, 2004) analyzed optimal commodity taxation in settings with non-linear income taxation and production inefficiencies. He showed that commodity tax differentiation is optimal even when preferences are identical and weakly separable. Our results clarify that this is not due to production inefficiencies as such. We showed that the Atkinson-Stiglitz theorem is recovered even if production decisions are inefficient. The results of Naito (1999, 2004) are therefore due to general-equilibrium effects on commodity prices, which are absent in our partial-equilibrium framework with constant commodity prices.

3.6 Optimal production inefficiency

The next proposition states the main result of this paper.

Proposition 2 *Aggregate production efficiency is not socially optimal. The optimal non-linear tax wedge on output of commodity x_i relative to the tax on the output of the numéraire commodity x_0 is given by:*

$$\begin{aligned}\frac{T_i'(p_i y_i(n)) - T_0'(y_0(n))}{(1 - T_i'(p_i y_i(n)))(1 - T_0'(y_0(n)))} \quad = \quad & -\frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)} \frac{\partial \ln(f_{y_i}(\cdot)/f_{y_0}(\cdot))}{\partial \ln n} \\ & + \frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)} \frac{n f_n(\cdot)}{l_0 f_{l_0}(\cdot)} \frac{\partial \ln(f_{y_i}(\cdot)/f_{y_0}(\cdot))}{\partial \ln l_0}, \quad \forall n, i. \end{aligned} \tag{39}$$

Proof. Use (33) for y_i and y_0:

$$\kappa f_{y_i} h(n) + \eta p_i h(n) + \theta u_{l_0} \left(\frac{f_{ny_i} f_{l_0} - f_{l_0 y_i} f_n}{f_{l_0}^2} \right) = 0, \tag{40}$$

$$\kappa f_{y_0} h(n) + \eta h(n) + \theta u_{l_0} \left(\frac{f_{ny_0} f_{l_0} - f_{l_0 y_0} f_n}{f_{l_0}^2} \right) = 0. \tag{41}$$

Subtract both equations to find:

$$\kappa (f_{y_i} - f_{y_0}) h(n) + \eta (p_i - 1) h(n) = -\frac{\theta u_{l_0}}{f_{l_0}^2} \left(f_{ny_i} f_{l_0} - f_{l_0 y_i} f_n - f_{ny_0} f_{l_0} + f_{l_0 y_0} f_n \right). \tag{42}$$

Use (8) to find $f_{y_i} = f_{y_0} p_i \frac{(1-T_i')}{(1-T_0')}$ and substitute in (42):

$$\frac{\kappa}{\eta} f_{y_0} \left(p_i \frac{(1-T_i')}{(1-T_0')} - 1 \right) + (p_i - 1) = -\frac{n u_{l_0}}{u_{x_0}} \frac{u_{x_0} \theta / \eta}{n h(n)} \frac{\left(f_{ny_i} f_{l_0} - f_{l_0 y_i} f_n - f_{ny_0} f_{l_0} + f_{l_0 y_0} f_n \right)}{f_{l_0}^2}. \tag{43}$$

Use (33) for y_0 to derive

$$\frac{\kappa}{\eta} f_{y_0} = -1 - \frac{n u_{l_0}}{u_{x_0}} \frac{u_{x_0} \theta / \eta}{n h(n)} \left(\frac{f_{ny_0} f_{l_0} - f_{l_0 y_0} f_n}{f_{l_0}^2} \right). \tag{44}$$

Substitution of (44) in (43) and rearranging gives:

$$p_i \left(\frac{T_i' - T_0'}{1 - T_0'} \right) = -\frac{n u_{l_0}}{u_{x_0}} \frac{u_{x_0} \theta / \eta}{n h(n)} \left[\frac{\left(f_{ny_i} f_{l_0} - f_{l_0 y_i} f_n \right)}{f_{l_0}^2} - \left(\frac{f_{ny_0} f_{l_0} - f_{l_0 y_0} f_n}{f_{l_0}^2} \right) p_i \frac{(1-T_i')}{(1-T_0')} \right]. \tag{45}$$

Use (10) for $k = 0$ to find $\frac{u_{l_0}}{u_{x_0}} = (1 - T_0') \frac{f_{l_0}}{f_{y_0}}$ and use (8) to find $\frac{f_{y_i}}{f_{y_0}} = p_i \frac{(1-T_i')}{(1-T_0')}$. Substitute both in (45) to find:

$$\frac{(T_i' - T_0')}{(1 - T_1')(1 - T_0')} = -\frac{n u_{x_0} \theta / \eta}{n h(n)} \left(\frac{f_{ny_i} f_{l_0} - f_{l_0 y_i} f_n}{f_{l_0} f_{y_i}} - \frac{f_{ny_0} f_{l_0} - f_{l_0 y_0} f_n}{f_{l_0} f_{y_0}} \right). \tag{46}$$

Derive that $\frac{1}{n} \left(\frac{f_{ny_i}}{f_{y_i}} - \frac{f_{ny_0}}{f_{y_0}} \right) = \frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n}$ and $\frac{1}{l_0} \left(\frac{f_{l_0 y_i}}{f_{y_i}} - \frac{f_{l_0 y_0}}{f_{y_0}} \right) = \frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln l_0}$ and rewrite so as to establish the proposition. ■

Recall from the section on first-best policies above that aggregate production efficiency is obtained when $f_{y_0}/f_{y_i} = p_i$. In that case, the marginal rates of transformation between x_i- and x_0-goods f_{y_0}/f_{y_i} for all commodities i are equalized for all individuals in the economy. From Proposition 2 follows that commodity outputs are generally taxed at differential rates, hence production inefficiency is generally desirable. To see why, we have $f_{y_0}/f_{y_i} - p_i > 0$ (< 0) if the output of x_i goods is taxed at a higher (lower) rate than output of the numéraire commodity x_0, i.e., $T_i' > T_0'$ ($T_i' < T_0'$), see first-order condition (8). Consequently, aggregate production decisions are generally not efficient (i.e., $f_{y_0}/f_{y_i} \neq p_i$). Production inefficiencies are optimal for two distinct reasons, which are the mirror image for the reasons why consumption inefficiencies are optimal.

First, if workers with a higher ability have a comparative advantage in producing the x_i-commodity over producing the numéraire x_0-commodity, then we have $f_{ny_i}/f_{y_i} < f_{ny_0}/f_{y_0}$(and vice versa). Consequently, the marginal rate of transformation decreases in ability as higher ability workers are able to produce *relatively* more output of x_i-commodities than the x_0-commodities with an additional unit of factor input: $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n} > 0$. Thus, when the marginal rate of transformation between x_i- and x_0-goods f_{y_i}/f_{y_0} decreases (increases) with ability n, output of x_i-commodities should be taxed at higher (lower) rates than output of x_0-commodities. Intuitively, when an individual of a higher ability is relatively more (less) productive in producing x_i-goods, this individual will allocate more (less) of his factor inputs to producing these goods. Consequently, the output from this production activity reveals information on the hidden ability of this individual, which should optimally be exploited for income redistribution.

We thus uncover a production counterpart of the results by Mirrlees (1976) and Saez (2002). These authors showed that if the marginal rates of *substitution* in consumption vary with ability, then differential commodity taxation is optimal to redistribute income, see also the previous section. Intuitively, the government wishes to tax commodities that the high-ability types like to consume. We, in contrast, demonstrated that when marginal rates of *transformation* vary with ability, the government should set higher taxes on the production of commodities in which the high ability types have a comparative advantage. The government thus wishes to tax outputs of commodities that high-ability individuals like to produce most.

Second, the term $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln l_0}$ captures to what extent labor supply is more complementary to the production of y_i commodities than to y_0 commodities. It captures how the marginal rate of transformation varies with labor supply. Outputs of commodities that are more complementary to labor should be taxed at lower rates, and vice versa. One can give a Ramsey-type intuition to these results. Suppose that outputs of all commodities are independent, so that $\frac{\partial f_{y_i}}{\partial y_j} = 0$, then $\left(\frac{\partial \ln f_{y_i}}{\partial \ln l_0}\right)^{-1}$ can be interpreted as an *implicit* labor-demand elasticity in the production of commodity x_i. If the implicit labor-demand elasticity in production of x_i commodities $(-\partial \ln f_{y_i}/\partial \ln l_i)^{-1}$ is higher relative to the implicit labor-demand elasticity in production of the numéraire commodity $(-\partial \ln f_{y_0}/\partial \ln l_0)^{-1}$, the output of commodity x_i should optimally be taxed less than the output of the numéraire commodity x_0. Intuitively, the government wishes to introduce production distortions to alleviate the distortions on total factor supply. By distorting the composition of production activities, the government induces individuals to allocate their labor time towards production sectors where taxing output gives fewer labor market distortions. Doing so reduces distortions in total factor supply $l_0(n)$ at the cost of distorting production activities.

These findings are the production counterpart of the Corlett and Hague (1953), Atkinson and Stiglitz (1976) and Jacobs and Boadway (2014) results for non-uniform commodity taxation. Commodities are not equally complementary to leisure when the marginal rates of *substitution* between commodities vary with labor effort. We demonstrate that a similar intuition applies to the production side of the economy since non-uniform output taxation is optimal: when the marginal rates of *transformation* between commodities vary with labor effort, outputs of commodities that are not equally complementary to labor and should be taxed at differential rates.

We can consider two special cases that further illustrate the two main reasons for introducing production distortions. First, suppose that the production technologies are given by $f(\mathbf{y}(n), \mathbf{l}(n), n) = \phi(\mathbf{y}(n), n) - \varphi(\mathbf{l}(n)) = 0$. In this case, we do allow for comparative advantage, since ability n does affect the production of all outputs differently, so that the marginal rates of transformation change with ability: $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n} = \frac{\partial \ln(\phi_{y_i}/\phi_{y_0})}{\partial n} \neq 0$. But, the marginal rates of transformation between any pair of outputs are constant in labor input: $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln l_0} = \frac{\partial \ln(\phi_{y_i}/\phi_{y_0})}{\partial \ln l_0} = 0$. Thus, the implicit labor demand elasticities are infinite, cf. $\frac{\partial \ln f_{y_i}}{\partial \ln l_0} = \frac{\partial \ln f_{y_0}}{\partial \ln l_0} = 0$. Consequently, production distortions are introduced only to exploit comparative advantages, but not to alleviate labor market distortions.

An example of this technology is the following. Suppose we have only one labor input l and two production outputs y_0 and y_1. Let the production set be described by a CES-technology:

$$f(y_0(n), y_1(n), l(n), n) \equiv \left[\frac{\gamma}{n^\alpha} (y_0(n))^\rho + \frac{(1-\gamma)}{n^\beta} (y_1(n))^\rho \right]^{\frac{1}{\rho}} - l(n) = 0, \quad \alpha, \beta, \rho > 0, \quad 0 < \gamma < 1. \tag{47}$$

The marginal rate of transformation between y_1 and y_0 goods is: $f_{y_1}/f_{y_0} = \frac{(1-\gamma)}{\gamma} n^{(\alpha-\beta)} \left(\frac{y_1}{y_0} \right)^{\rho-1}$. Consequently, we have $\frac{\partial \ln(f_{y_1}/f_{y_0})}{\partial \ln l_0} = 0$ and $\frac{\partial \ln(f_{y_1}/f_{y_0})}{\partial \ln n} = \alpha - \beta > 0$. In this case, an individual with a high ability n has a comparative advantage in producing y_0 commodities over y_1 commodities if $\alpha < \beta$. For the same labor time allocated to production, the individual produces more y_0 goods than y_1 goods. Consequently, output of y_1-commodities should be taxed at a higher rate.

Second, suppose now that comparative advantage is absent and production is described by the following technology: $f(\mathbf{y}(n), \mathbf{l}(n), n) = \varphi(n)\phi(\mathbf{y}(n), \mathbf{l}(n)) = 0$. In this case, we no longer have comparative advantage, since ability affects all inputs and outputs symmetrically: $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n} = \frac{\partial \ln(\phi_{y_i}/\phi_{y_0})}{\partial \ln n} = 0$. However, the marginal rates of transformation are not constant in labor effort: $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln l_0} = \frac{\partial \ln(\phi_{y_i}/\phi_{y_0})}{\partial \ln l_0} \neq 0$. Production distortions are thus introduced to alleviate labor market distortions. An example is the following. Suppose we have only one labor input and two production outputs. The production set is described by the following technology:

$$f(y_0(n), y_1(n), l(n), n) \equiv n \left(\frac{y_0(n)}{(l(n))^\alpha} + \frac{y_1(n)}{(l(n))^\beta} - 1 \right) = 0, \quad \alpha, \beta > 0. \tag{48}$$

In this case labor effort produces more y_0 goods than y_1 goods if $\alpha > \beta$. The marginal rate of transformation is given by $\frac{f_{y_1}}{f_{y_0}} = l^{\alpha-\beta}$. Consequently, we have $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n} = 0$ and $\frac{\partial \ln(f_{y_1}/f_{y_0})}{\partial \ln l} = \alpha - \beta$. Thus, good y_0 should be taxed at a lower rate than good y_1 to reduce labor market distortions.

How do our findings relate to Diamond and Mirrlees (1971)? Aggregate production efficiency optimal in their analysis because all individuals have access to identical, constant returns-to-scale production technologies. We can recover the production efficiency theorem if individuals would differ only in their endowments of efficiency units of labor, and all production activities would entail identical production technologies.

Corollary 1 *If the production sets are identical for all individuals and given by*

$$f(\mathbf{y}(n), \mathbf{l}(n), n) \equiv \mathbf{A}'\mathbf{y}(n) - n\mathbf{l}(n) = 0, \quad \forall n. \tag{49}$$

where $\mathbf{A} \equiv (A_0, A_1, \cdots, A_I)$ *is a common vector of technological coefficients, then aggregate production efficiency is obtained*

$$\frac{f_{y_0}(\mathbf{y}(n), \mathbf{l}(n), n)}{f_{y_i}(\mathbf{y}(n), \mathbf{l}(n), n)} = \frac{A_0}{A_i} = p_i, \quad \forall n, i, \tag{50}$$

and all outputs are taxed at equal rates $T_i'(p_i y_i(n)) = T_0'(y_0(n))$.

In this particular case, marginal rates of transformation are constant and equal to $f_{y_0}/f_{y_i} = A_0/A_i$. The marginal rates of transformation do neither vary with ability n nor with labor l_0. Hence, there is no comparative advantage $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln n} = 0$ for all i and n. And, implicit labor-demand elasticities are infinitely elastic in all production activities, since $\frac{\partial \ln(f_{y_i}/f_{y_0})}{\partial \ln l_0} = 0$. In this case, there are neither distributional nor efficiency reasons to distort production activities, and we find aggregate production efficiency as the social optimum in second best. The corollary assumes that the $\mathbf{A}$-vector is constant, which is necessary in our partial-equilibrium setup so as to ensure constant prices. However, the results would generalize to settings where the marginal rates of transformation are determined endogenously in general equilibrium, but would be the same for all individuals. This follows from the next point.

From the incentive compatibility constraints we can reveal a second intuition behind the desirability of aggregate production inefficiency in Diamond and Mirrlees (1971). Note that if $f(\mathbf{y}(n), \mathbf{l}(n), n) = \mathbf{A}'\mathbf{y}(n) - n\mathbf{l}(n) = 0$, the incentive-compatibility constraint (22) can be rewritten as:

$$\frac{\mathrm{d}u(n)}{\mathrm{d}n} = -u_{l_0}(\mathbf{x}(n), \mathbf{l}(n), n)\frac{\mathbf{l}(n)}{n} + u_n(\mathbf{x}(n), \mathbf{l}(n), n), \quad \forall n. \tag{51}$$

In this case the incentive-compatibility constraint (22) is the same as in Mirrlees (1971) if preferences would be identical and $u_n = 0$. Thus, when production technologies of all individuals are equal, then the allocation of outputs of various commodities $\mathbf{y}(n)$ does not affect the incentive-compatibility constraints. Intuitively, individuals of a higher ability type cannot mimic individuals of a lower ability by reallocating their labor input to produce more of certain outputs, as their labor productivity per hour worked is the same in the production of all commodities. Naturally, production decisions should be overall efficient then. This sheds a different light on the meaning of the production efficiency theorem. As long as every individual can freely trade all commodities, while everyone has the same access to the aggregate production technology, it is not optimal to distort production decisions for income redistribution. Marginal rates of transformation no longer reveal any information on ability.

Corollary 2 generalizes the special case of Diamond-Mirrlees to the case where there is a Cobb-Douglas production set with uniform elasticities α.

Corollary 2 *If production sets are Cobb-Douglas, i.e.,*

$$f(\mathbf{y}(n), \mathbf{l}(n), n) \equiv \prod_{i=1}^{I} A_i(y_i(n))^{\alpha} - n\phi(\mathbf{l}(n)) = 0, \quad \alpha \leq 1/I, \tag{52}$$

Then aggregate production efficiency is obtained:

$$\frac{f_{y_0}(\mathbf{y}(n),\mathbf{l}(n),n)}{f_{y_i}(\mathbf{y}(n),\mathbf{l}(n),n)}=\frac{A_0}{A_i}\left(\frac{y_0(n)}{y_i(n)}\right)^{\alpha-1}=p_i,\quad \forall n,i, \tag{53}$$

In this case high-ability individuals have no longer a comparative advantage in the production of any commodity, since $\frac{\partial\ln(f_{y_i}/f_{y_0})}{\partial\ln n}=0$. We also have that the implicit labor-demand elasticities are equalized across all activities: $\frac{\partial\ln(f_{y_i}/f_{y_0})}{\partial\ln l_0}=0$. Thus, there would be neither redistributive reasons – exploiting comparative advantage – nor efficiency reasons – alleviating labor supply distortions – to employ differential output taxes. This special case is interesting because it demonstrates that it is not decreasing returns to scale in production as such – giving rise to quasi-rents – that generates the violation of production efficiency, see for example Diamond and Mirrlees (1971), Stiglitz and Dasgupta (1971) and Mirrlees (1972). Our tax code is complete and there are no untaxed pure rents, see also the household budget constraint (3). Indeed, production distortions are desirable only to the extent that they help to redistribute *ability-specific* rents. If the elasticities α would differ across production activities i, then production inefficiencies become optimal, as was shown above in a special case.

The Diamond and Mirrlees (1971) and Cobb-Douglas cases are not the only examples where production efficiency is optimal. The critical question is under which properties on the production technology production efficiency is obtained. These conditions are, again, the mirror conditions of the Atkinson-Stiglitz theorem for consumption efficiency as the next proposition shows.

Proposition 3 *Aggregate production efficiency is obtained when the individual production sets $f(\cdot)$ are weakly separable between outputs $\mathbf{y}(n)$ on the one hand and inputs $\mathbf{l}(n)$ and ability n, on the other hand:*

$$f(\mathbf{y}(n),\mathbf{l}(n),n)\equiv g(\phi(\mathbf{y}(n)),\mathbf{l}(n),n)=0, \tag{54}$$

where $\phi(\mathbf{y}(n))$ is a sub-production function that is independent from n. The aggregate production set can then be written as:

$$\int_{\mathcal{N}} f(\mathbf{y}(n),\mathbf{l}(n),n)\mathrm{d}H(n)=\int_{\mathcal{N}} g(\phi(\mathbf{y}(n)),\mathbf{l}(n),n)\mathrm{d}H(n)=0. \tag{55}$$

Aggregate production efficiency does not require that an aggregate production technology, which is defined only over aggregate inputs and aggregate outputs, exists. However, reverse is always true. To see this, let the production technology be written as a function of the aggregate outputs and inputs:

$$\int_{\mathcal{N}} f(\mathbf{y}(n),\mathbf{l}(n),n)\mathrm{d}H(n)\equiv F\left(\int_{\mathcal{N}}\mathbf{y}(n)\mathrm{d}H(n),\int_{\mathcal{N}}\mathbf{l}(n)\mathrm{d}H(n)\right)=0. \tag{56}$$

In this case, aggregate production efficiency results trivially, since the marginal rates of transformation between any pair of commodities are independent from ability type n and individual labor supply $\mathbf{l}(n)$, provided that each individual is infinitely small and cannot influence aggregate economic conditions.

Only under specific conditions individual production technologies can be aggregated into one aggregate production technology.[8] Gorman (1996) shows that as long as individual production technologies belong to the Gorman polar form such aggregation is feasible as long as prices are linear. This implies that supplies of outputs or demands for inputs in production feature linear Engel curves in gross profits (value of output minus costs of production). These conditions are not only very strict, but also rule out non-linear income taxation. Again, these conditions mirror the conditions under which it is not useful to have linear commodity tax differentiation, see Sandmo (1974), Atkinson and Stiglitz (1976) and Deaton (1979).

The production efficiency theorem is applicable to any subset of commodities that are produced with identical technologies, which can be aggregated into an aggregate production function. However, production efficiency would not be optimal for the remaining commodities, where individuals do not share the same technological possibilities to produce them. Similarly, we could allow for different firms, each having different production technologies, in which subsets of the population work. This can be modelled by having identical individual production functions for each subset of the population working in the same firm. Then, there will be 'production efficiency' within these subsets of the population, but not across the subsets of the population.

Our interpretation is also related to the findings of Saez (2004). He analyzes optimal income and commodity taxation in models in which individuals make an occupational choice. In his analysis, the relative supply of labor for each occupation is infinitely elastic. Since all individuals can supply labor in all occupations, perfect arbitrage on the labor market removes all factor price differentials per efficiency unit of labor. All marginal rates of transformation become identical across individuals. In other words: perfect arbitrage ensures that the production technology becomes the same for all individuals. The only reason why individuals differ is their endowments of efficiency units of labor. Hence, the standard Diamond-Mirrlees and Atkinson-Stiglitz theorems apply given the usual conditions of weak separability, constant returns to scale, etc.

Finally, our analysis reveals that optimal output taxes generally depend on parameters from the production side of the model. Hence, the property in Diamond and Mirrlees (1971) that optimal tax formulae only depend on the parameters from the consumption side of the economy – dubbed the 'Tax Formula result' in Saez (2004) – is no longer applicable.

3.7 Optimal income taxation

Finally, we derive the optimal non-linear income tax on the output of the numéraire good, which we will refer to as the non-linear income tax.

Proposition 4 *The optimal non-linear income tax schedule on production of x_0-commodities*

[8] This question goes back to the old Cambridge-Cambridge controversy on whether an aggregate production function can exist. See also Cohen and Harcourt (2003) for an overview of that debate.

is given by:

$$
\begin{aligned}
\frac{T_0'(y_0(n))}{1-T_0'(y_0(n))} &= \frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)}\frac{nf_n}{l_0 f_{l_0}}\left(\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0}-\frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0}\right) \\
&\quad +\frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)}\left(-\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n}+\frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln n}\right), \quad \forall n,
\end{aligned} \tag{57}
$$

$$
\begin{aligned}
\frac{\theta(n)}{\eta} &= \int_n^{\overline{n}}\left(\frac{1}{u_{x_0}(\cdot)}-\frac{1}{\eta}\right) \\
&\quad \times \exp\left[\int_n^m \left(\frac{\partial \ln u_{x_0}}{\partial \ln s}-\frac{sf_s}{l_0 f_{l_0}}\frac{\partial \ln u_{x_0}}{\partial \ln \ell}\right)\frac{\mathrm{d}s}{s}\right] h(m)\mathrm{d}m > 0 \quad \forall n, \neq \underline{n}, \overline{n}.
\end{aligned} \tag{58}
$$

Proof. We can rewrite first-order condition (34) for l_0 using the derivatives of x_0 (30) and the first-order condition (10) – omitting the indices and functional arguments where we evaluate the expression only for l_0:

$$
\begin{aligned}
\frac{\kappa}{\eta}f_{y_0}+(1-T_0') &= -\frac{f_{y_0}}{f_{l_0}}\frac{\theta/\eta}{h(n)}\frac{f_n}{f_{l_0}}\left(u_{l_0 l_0}-u_{l_0 x_0}\frac{u_{l_0}}{u_{x_0}}\right) \\
&\quad -\frac{f_{y_0}}{f_{l_0}}\frac{\theta/\eta}{h(n)}\left[u_{l_0}\left(\frac{f_{nl_0}f_{l_0}-f_{l_0 l_0}f_n}{f_{l_0}^2}\right)-\left(u_{nl_0}-u_{nx_0}\frac{u_{l_0}}{u_{x_0}}\right)\right].
\end{aligned} \tag{59}
$$

From the first-order condition for y_0 (33) follows

$$
\frac{\kappa}{\eta}f_{y_0} = -1-\frac{nu_{l_0}}{u_{x_0}}\frac{u_{x_0}\theta/\eta}{nh(n)}\left(\frac{f_{ny_0}f_{l_0}-f_{l_0 y_0}f_n}{f_{l_0}^2}\right). \tag{60}
$$

Substitution of (60) in (59) gives:

$$
\begin{aligned}
T_0' &= \frac{f_{y_0}}{f_{l_0}}\frac{\theta/\eta}{h(n)}\left[\frac{f_n}{f_{l_0}}\left(u_{l_0 l_0}-u_{l_0 x_0}\frac{u_{l_0}}{u_{x_0}}\right)+u_{l_0}\left(\frac{f_{nl_0}f_{l_0}-f_{l_0 l_0}f_n}{f_{l_0}^2}\right)\right] \\
&\quad -\frac{f_{y_0}}{f_{l_0}}\frac{\theta/\eta}{h(n)}\left(u_{nl_0}-u_{nx_0}\frac{u_{l_0}}{u_{x_0}}\right)-\frac{nu_{l_0}}{u_{x_0}}\frac{u_{x_0}\theta/\eta}{nh(n)}\left(\frac{f_{ny_0}f_{l_0}-f_{l_0 y_0}f_n}{f_{l_0}^2}\right).
\end{aligned} \tag{61}
$$

Use the first-order condition for y_0 in (8) to find $\frac{u_{l_0}}{u_{x_0}} = (1-T_0')\frac{f_{l_0}}{f_{y_0}}$, substitute this in (61) and rearrange:

$$
\begin{aligned}
\frac{T_0'}{1-T_0'} &= \frac{u_{x_0}\theta/\eta}{nh(n)}\left[\frac{n}{u_{l_0}}\frac{f_n}{f_{l_0}}\left(u_{l_0 l_0}-u_{l_0 x_0}\frac{u_{l_0}}{u_{x_0}}\right)+n\left(\frac{f_{nl_0}f_{l_0}-f_{l_0 l_0}f_n}{f_{l_0}^2}\right)\right] \\
&\quad -\frac{u_{x_0}\theta/\eta}{nh(n)}\left[\frac{n}{u_{l_0}}\left(u_{nl_0}-u_{nx_0}\frac{u_{l_0}}{u_{x_0}}\right)+\frac{nf_{l_0}}{f_{y_0}}\left(\frac{f_{ny_0}f_{l_0}-f_{l_0 y_0}f_n}{f_{l_0}^2}\right)\right].
\end{aligned} \tag{62}
$$

Note that $\frac{l_0 u_{l_0 l_0}}{u_{l_0}} - \frac{l_0 u_{l_0 x_0}}{u_{x_0}} = \frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0}$, $\frac{n u_{n l_0}}{u_{l_0}} - \frac{n u_{n x_0}}{u_{x_0}} = \frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n}$, $\frac{n f_{n l_0}}{f_{l_0}} - \frac{n f_{n y_0}}{f_{y_0}} = \frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n}$ and $\frac{l_0 f_{l_0 l_0}}{f_{l_0}} - \frac{l_0 f_{l_0 y_0}}{f_{y_0}} = \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0}$. Hence, we find the first part of the Proposition:

$$\begin{aligned} \frac{T_0'}{1-T_0'} &= \frac{u_{x_0}\theta/\eta}{nh(n)} \frac{n f_n}{l_0 f_{l_0}} \left(\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0} - \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0} \right) \\ &\quad - \frac{u_{x_0}\theta/\eta}{nh(n)} \left(\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n} - \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln n} \right). \end{aligned} \tag{63}$$

Further, use the derivatives in equation (30) in the first-order condition for utility (35) to get:

$$\frac{\mathrm{d}\theta}{\mathrm{d}n} + \theta \left(\frac{u_{n x_0}}{u_{x_0}} - \frac{f_n}{f_{l_0}} \frac{u_{l_0 x_0}}{u_{x_0}} \right) = \left(1 - \frac{\eta}{u_{x_0}} \right) h(n). \tag{64}$$

Note that $\frac{n u_{n x_0}}{u_{x_0}} = \frac{\partial \ln u_{x_0}}{\partial \ln n}$ and $\frac{l_0 u_{l_0 x_0}}{u_{x_0}} = \frac{\partial \ln u_{x_0}}{\partial \ln \ell}$:

$$\frac{\mathrm{d}\theta}{\mathrm{d}n} + \frac{\theta}{n} \left(\frac{\partial \ln u_{x_0}}{\partial \ln n} - \frac{n f_n}{l_0 f_{l_0}} \frac{\partial \ln u_{x_0}}{\partial \ln \ell} \right) = \left(1 - \frac{\eta}{u_{x_0}} \right) h(n). \tag{65}$$

This is a linear differential equation in θ of the form $\frac{\mathrm{d}\theta(n)}{\mathrm{d}n} + a(n)\theta(n) = b(n)$, with $a(n) \equiv \frac{1}{n}\left(\frac{\partial \ln u_{x_0}}{\partial \ln n} - \frac{n f_n}{l_0 f_{l_0}} \frac{\partial \ln u_{x_0}}{\partial \ln \ell} \right)$ and $b(n) \equiv (1 - \frac{\eta}{u_{x_0}(\cdot)})h(n)$. This differential equation can be integrated, using a transversality condition from (36) to find: $\theta(n) = -\int_n^{\overline{n}} \exp\left[\int_n^m a(s)\mathrm{d}s\right] b(m)dm$. Substituting for $a(n)$ and $b(n)$ yields the second part of the proposition. ■

The major difference of our expression for the optimal non-linear tax in (57) with the one derived by Mirrlees (1971) is the elasticity term $\frac{n f_n}{l_0 f_{l_0}} \left(\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0} - \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0} \right) - \frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n} + \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln n}$. This term consists of four elements. First, it encompasses the distortions of the income tax on labor supply as captured by $\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0}$, which equals the inverse of the standard Frisch elasticity of labor supply. This term is the main ingredient of the elasticity of the tax base in Mirrlees (1971). The higher is the elasticity of labor supply, the lower should be the optimal tax rate. Second, the production distortion follows from $\frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0}$. When the tax on the output of the numéraire commodity x_0 is increased, the individual will allocate less labor to the production of the numéraire commodity, and more to the production of other commodities. This effect is new compared to Mirrlees (1971), since that analysis only considers one production sector. In our setting there is a correction $\frac{n f_n}{l_0 f_{l_0}}$ reflecting the fact that outputs and earnings are determined by individual production functions. Third, $\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n}$ captures how the willingness to supply labor varies with the skill level – conditional on labor income – since we allowed for preference heterogeneity. It measures how ability affects the elasticity of labor supply. Fourth, $\frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln n}$ captures how the production elasticities change with ability. All four elements determine the effective elasticity of the total tax base. Note again that the elasticity term generally depends on parameters on the production side of the economy.

The social marginal value of income redistribution $\theta(n)/\eta$ at skill level n in equation (59) is the same as in Mirrlees (1971), except for the presence of $\frac{\partial \ln u_{x_0}(\cdot)}{\partial \ln n}$ in the bracketed exp[·]-term inside the integral. $\frac{\partial \ln u_{x_0}(\cdot)}{\partial \ln n}$, again, originates from the fact that we allowed for preference heterogeneity. If the utility function would be the same for all individuals, it would disappear. Saez (2001) and Jacobs and Boadway (2014) show that the term in brackets is associated with income effects in labor effort.

The expression for the optimal non-linear income tax is otherwise very similar to the expression found in Mirrlees (1971). To see this, suppose that the assumptions of the Diamond and Mirrlees (1971) production efficiency theorem would hold and there would be no preference heterogeneity. In particular, if individual production technologies are given by $f(\mathbf{y}(n), \mathbf{l}(n), n) \equiv \mathbf{A}'\mathbf{y}(n) - n\mathbf{l}(n) = 0$, and the utility function is $u(\mathbf{x}, \mathbf{l}, n) = u(\mathbf{x}, \mathbf{l})$, then we can derive that $\frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln n} = \frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln l_0} = 0$, $\frac{\partial \ln(f_{l_0}/f_{y_0})}{\partial \ln n} = 1$, and $\frac{n f_n}{l_0 f_{l_0}} = 1$. The optimal non-linear income tax would be given by:

$$\frac{T_0'(y_0(n))}{1 - T_0'(y_0(n))} = \frac{u_{x_0}(\cdot)\theta(n)/\eta}{nh(n)} \left(1 + \frac{\partial \ln(u_{l_0}/u_{x_0})}{\partial \ln l_0}\right) \tag{66}$$

Which is exactly the same expression as in Mirrlees (1971). Note that under these assumptions, there would be overall production efficiency, hence the tax schedule is the same for all outputs x_i: $T_i'(p_i y_i(n)) = T_0'(y_0(n))$. The interested reader may wish to consult Diamond (1998) and Saez (2001) for more interpretations and intuitions of the non-linear tax schedule.

4 Policy implications

When not everyone has the same technological possibilities to transform inputs into outputs, the practical desirability of free trade, no taxation of intermediate goods, the use of market prices in social-cost benefit analysis or public sector production, might all be called into question. Our analysis has therefore a number of potentially important policy-relevant implications, which we will shortly discuss.

4.1 Capital income taxation

Our model could be given an intertemporal interpretation, with commodity x_0 denoting consumption today and commodities x_i consumption levels at future dates. In such a context, the marginal rates of transformation f_{y_0}/f_{y_i} can be viewed as the technological opportunities to transform current consumption into future consumption. These production functions may differ by individuals for various reasons. First, high-ability individuals might generate larger returns on their savings than low-ability individuals do. For example, when they earn returns on assets in their own, closely-held firms, which are also determined by individual productive abilities, see for example Gerritsen et al. (2015). Second, more able individuals might earn larger returns on their assets due scale effects in portfolio management (Piketty (2014)). Third, some individuals might be barred from entering capital markets. If high-ability individuals are less liquidity constrained than low-ability individuals are, then they are typically more efficient to transform current consumption into future consumption. Fourth, when insurance markets are missing, different individuals have different possibilities to transform current consumption into *expected* future consumption. If risk risk aversion falls with productive ability, the consumption-possibilities frontier in expected consumption will be less concave. For all these reasons, the marginal rates of transformation in consumption differ by individual's abilities. Consequently, taxes on saving can be socially desirable for redistributional reasons, as Gerritsen et al. (2015)

formally demonstrate.

4.2 Intermediate goods taxation

Jacobs and Bovenberg (2011) develop a completely worked out example where taxes or subsidies on intermediate goods taxation are socially desirable. In their model, educational investment is an intermediate good used in human capital production. Human capital is employed in final goods production. Education should be taxed if high-ability individuals have a comparative advantage in human capital formation. Education should be subsidized if this alleviates labor market distortions. Whether human capital should be taxed or subsidized on a net basis remains ambiguous as it is the result of these two off-setting effects. Our analysis shows that taxes or subsidies on the use of intermediate goods can be optimal more generally. In our model intermediate goods may be seen as negative x_i's in the production technology. Intermediate goods should be taxed (subsidized) if high-ability (low-ability) individuals have a comparative advantage in using these intermediate goods in final-goods production. Moreover, intermediate goods should be taxed less (more) if they are stronger (weaker) complements with labor supply.

4.3 Differential sector taxation

Our results demonstrated that different outputs should be taxed at different rates so that the playing field between sectors/occupations should not be level. Outputs produced in sectors/occupations in which high-ability individuals have a comparative advantage – for example IT of finance – should be taxed at higher rates than the outputs from sectors/occupations in which low-ability individuals have a comparative advantage – for example restaurants. Similarly, our findings might also explain why optimal tax rates should be lower in sectors/occupations in which labor demand is relatively more elastic. Examples include the construction sector, bars and restaurants or personal services. Due to the presence of close substitutes (household production, black/grey labor market) labor demand might be relatively more elastic in these sectors compared to other sectors/occupations.

4.4 Free trade

Our analysis showed that it is desirable to tax the output in which high-ability agents have a comparative advantage, or to subsidize the outputs those sectors in which low-ability individuals have a comparative advantage. This implies that in order to redistribute income in the most efficient way, trade tariffs, production subsidies, and so on, could be socially desirable to raise the net incomes of low-skilled workers. Indeed, some of the commodities in our model can be traded goods. Similarly, deviations from residence-based factor taxes or destination-based consumption taxes might be optimal. On the other hand, free trade could also make labor demand in certain sectors more elastic as skilled labor or capital may move more easily abroad. In that case, these sectors in which these inputs are more heavily used may be taxed less.

4.5 Public production and social cost-benefit analysis

The findings of this paper imply that it may be not be desirable to use market prices in social cost-benefit analysis. Indeed, the government may oversupply (undersupply) public goods if low-skilled (high-skilled) workers are have a comparative advantage in producing them. Similarly, the government may produce more public goods than dictated by conventional social cost-benefit analysis, if these public goods help to lower labor market distortions in the private sector, for example, through child-care facilities. Finally, there appears to be no obvious candidate for the correct social discount rate if individuals have different intertemporal marginal rates of transformation, see also our the discussion above.

5 Conclusions

The Diamond and Mirrlees (1971) production efficiency theorem is derived under the assumption that all individuals have access to the same technological possibilities to transform their inputs into outputs. These technological possibilities are described by the aggregate production function. Although an aggregate production technology is a useful device to describe many simple market transactions of homogeneous commodities or publicly traded assets, it does not seem plausible that all individuals have identical access to the same technological opportunities to transform their labor, assets and other resources into outputs, which can either be sold on the market, consumed, or saved. The main message of this paper is that when individuals face different technological possibilities, the production efficiency theorem generally breaks down.

We have shown that aggregate production efficiency is not desirable for either equity or efficiency reasons. Production of goods should be taxed at higher (lower) rates if individuals with higher (lower) earnings abilities have a comparative advantage in the production of these goods. Outputs of certain goods should also be taxed at higher rates if labor supply is less complementary with these outputs. When production technologies feature weak separability of outputs from labor, and production of outputs are independent from ability, then the production efficiency theorem is obtained.

For future research it is important to empirically examine to what extent individuals indeed operate different production technologies. This is not an easy task. Individual's factor incomes (labor and capital incomes) are the result of many forces such as individual productive abilities, occupational and human capital decisions, access to markets (for labor, capital and insurance), general-equilibrium effects on factor prices, and so on.

If not all individuals have access to the same technological opportunities, then many of the very strong policy prescriptions that follow from Diamond and Mirrlees (1971) – free trade, no intermediate goods taxation, no sectoral differentiation in taxation, use of market prices in public production and social cost-benefit analysis – need not be applicable. Many economic policies that appear to be distortionary at first sight could turn out to be socially desirable after all.

References

Acemoglu, Daron, Aleh Tsyvinski, and Mikhail Golosov. 2008. "Political Economy of Intermediate Goods Taxation." *Journal of the European Economic Association* 6 (2/3):353–366.

Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1-2):55–75.

Bovenberg, A. Lans and Bas Jacobs. 2005. "Redistribution and Education Subsidies are Siamese Twins." *Journal of Public Economics* 89 (11-12):2005–2035.

Cohen, Avi J. and G. C. Harcourt. 2003. "Retrospectives: Whatever Happened to the Cambridge Capital Theory Controversies?" *Journal of Economic Perspectives* 17 (1):199–214.

Corlett, Wilfred J. and Douglas C. Hague. 1953. "Complementarity and the Excess Burden of Taxation." *Review of Economic Studies* 21 (1):21–30.

Deaton, Angus. 1979. "Optimally Uniform Commodity Taxes." *Economics Letters* 2 (4):357–361.

Diamond, Peter A. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1):83–95.

Diamond, Peter A. and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1):8–27.

Gaube, Thomas. 2005. "Income Taxation, Endogenous Factor Prices and Production Efficiency." *Scandinavian Journal of Economics* 107 (2):335–352.

Gerritsen, Aart, Bas Jacobs, Alexandra Rusu, and Kevin Spiritus. 2015. "Optimal Taxation of Capital Income when Individuals have Different Returns." mimeo, Max Planck Institute Munich/Erasmus University Rotterdam/KU Leuven.

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan. 2014. "Differential Taxation and Occupational Choice." mimeo, Toulouse School of Economics/Northwestern University.

Gorman, William M. 1996. "Measuring the Quantities of Fixed Factors." chap. 19. Oxford University Press, 325–352.

Guesnerie, Roger and Jesus K. Seade. 1982. "Nonlinear Pricing in a Finite Economy." *Journal of Public Economics* 17 (2):157–179.

Jacobs, Bas and Robin Boadway. 2014. "Optimal Linear Commodity Taxation under Optimal Non-Linear Income Taxation." *Journal of Public Economics* 117 (1):201–210.

Jacobs, Bas and A. Lans Bovenberg. 2011. "Optimal Taxation of Human Capital and the Earnings Function." *Journal of Public Economic Theory* 13 (6):957–971.

Keen, Michael and David Wildasin. 2004. "Pareto-Efficient International Taxation." *American Economic Review* 94 (1):259–275.

Kleven, Henrik Jacobsen, Richter Wolfram F., Sørensen, and Peter Birch. 2000. "Optimal Taxation with Household Production." *Oxford Economic Papers* 52 (3):584–594.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2):175–208.

———. 1972. "On Producer Taxation." *Review of Economic Studies* 39 (1):105–111.

———. 1976. "Optimal Tax Theory: A Synthesis." *Journal of Public Economics* 6 (4):327–358.

Naito, Hisahiro. 1999. "Re-examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implication for Production Efficiency." *Journal of Public Economics* 71 (2):165–188.

———. 2004. "Endogenous Human Capital Accumulation, Comparative Advantage and Direct vs. Indirect Redistribution." *Journal of Public Economics* 88 (12):2685–2710.

Newberry, David M. 1986. "On the Desirability of Input Taxes." *Economics Letters* 20 (3):267–270.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Cambridge-MA: Harvard University Press.

Renes, Sander S. and Floris Zoutman. 2014. "Extending the Atkinson-Stiglitz Theorem: Implementing Second-Best Allocations Through the Simplest Tax Schedule." mimeo, University of Mannheim/Norwegian School of Economics.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1):205–229.

———. 2002. "The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics* 83 (2):217–230.

———. 2004. "Direct or Indirect Instruments for Redistribution: Short-Run versus Long-Run." *Journal of Public Economics* 88:503–518.

Sandmo, Agnar. 1974. "A Note on the Structure of Optimal Taxation." *American Economic Review* 64 (4):701–706.

Stiglitz, Joseph E. 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics* 17 (2):213–240.

Stiglitz, Joseph E. and Partha Dasgupta. 1971. "Differential Taxation, Public Goods and Economic Efficiency." *Review of Economic Studies* 38 (2):151–174.