Re-examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and Its Implication for Production Efficiency

Hisahiro Naito *
Department of Economics
The University of Michigan

First version April 15, 1996
Revised version August 27, 1997

Abstract

This paper shows that the result of "uniform commodity tax under non-linear income taxation" by Atkinson and Stiglitz (1976) no longer holds if the production side of an economy is taken into the consideration. In particular, imposing a non-uniform commodity tax can Pareto-improve welfare even under non-linear income taxation.

In addition, the paper shows that if the government uses a non-linear income tax system for income redistribution, then the introduction of distortions in the public sector can Pareto-improve welfare, contrary to the results in Diamond and Mirrlees (1971).

Keywords: Optimal Non-linear Income Tax, Optimal commodity tax, Production Efficiency

JEL classification: H21, H23

*phone: 1-313-996-8179; email: hisahiro@econ.lsa.umich.edu

1 Introduction

One of the important topics in the public finance literature is the relationship between income redistribution and economic efficiency. For example, many economists have been concerned with whether there should be intervention in commodity markets when the government is concerned with income redistribution. Using the seminal work of Mirrlees (1971), Atkinson and Stiglitz (1976) showed that if the government can use a non-linear income tax to redistribute income, and if the utility function is weakly separable between goods and leisure, then optimal commodity taxes should be uniform. In the public finance literature, this "uniform commodity taxation" result has been confirmed and cited in many papers and textbooks: Atkinson and Stiglitz (1980, pp 437), Laffont and Tirole (1993, pp 195), Christiansen (1984) and Konishi (1995).

This result has important implications for public policy. Combined with the result of "production efficiency" by Diamond and Mirrlees (1971), the result of "uniform commodity taxation" suggests that the government should keep the commodity markets efficient even if it is concerned with income redistribution and that a progressive tax system (a non-linear income tax system) in the labor market is sufficient for income redistribution. For example, the above two results imply that small open countries should not distort trade patterns.

In reality, however, there are many public policies which distort commodity markets, such as tariffs and capital income taxes as well as commodity taxes and subsidies. One natural question about those policies is why the government still uses them, contrary to the forecasts from the theory.[1]

In this paper, we re-examine the result of "uniform commodity taxation" and find that the result critically depends on assumptions of constant marginal costs of production. We will show that imposing a non-uniform commodity tax can Pareto-improve welfare, even when the government is using a Pareto-efficient, non-linear income tax system under weak separability of workers' utility functions once the assumption of constant marginal cost of production is abandoned and the production side

[1] One possible answer is that the utility function is not weakly separable. However, as argued by Atkinson and Stiglitz(1976), most econometric studies on demand functions find it difficult to reject weak separability of utility functions.

of the economy is explicitly introduced in the analysis.

The reason non-uniform commodity taxes can Pareto-improve welfare when the production side is considered is that distortions in the commodity markets can change factor prices, and a change in factor prices can relax the incentive problem of income redistribution.

To illustrate, consider a situation where the government is redistributing income from skilled workers to unskilled workers by a Pareto-efficient non-linear income tax system without a commodity tax. With such redistribution, the incentive compatibility constraint of skilled workers is binding, and the labor supply of unskilled workers is at less than an efficient level. In this situation, if the marginal cost of production is not constant, then the government can increase the wage of unskilled workers and decrease the wage of skilled workers by suitable deviations from uniform commodity taxes.[2] At the margin, such redistribution through the change of wages has a first-order effect on welfare, whereas the distortionary cost of a commodity tax is zero to first order. Therefore, deviations from uniform commodity taxes can Pareto-improve welfare.

The results also suggests that public enterprises should minimize cost based on implicit factor prices that differ from the market prices, contrary to the result in Diamond and Mirrlees (1971).[3] In particular, this paper shows that introducing distortions in the public sector can Pareto-improve welfare when the government uses a non-linear income tax system for income redistribution. The basic intuition for this result can be explained as follows: Consider a situation where the government is using Pareto-efficient non-linear income taxation under efficient public production. In this situation, if the government tells public enterprises to hire more unskilled labor than the efficient level, then the government can increase the relative wage ratio of unskilled labor. As a result, the government can indirectly redistribute income from the skilled to the unskilled and mitigate the incentive problem

[2]The mechanism of a commodity tax to change the factor prices when the marginal cost of production is not constant is essentially the same as the analysis of the tax incidence in a general equilibrium framework, such as that by Harberger (1962). In a standard two goods and two factors general equilibrium model of tax incidence, if the government imposes a commodity tax on skilled labor-intensive goods, then the wage of skilled labor decreases and the wage of unskilled labor increases. See footnote 14.

[3]Diamond and Mirrlees (1971) showed that the marginal rate of transformation (MRT) between inputs in the public sector should be equal to MRT in the private sector and production in the public sector should be efficient as long as the government control all commodity prices and the technology is constant returns to scale.

of redistributive income taxation. Again, the effects on the incentive problem have a first-order effect on welfare while the input distortion in public enterprises is of second-order importance. Thus, introduction of distortions in the public sector Pareto-improves welfare.

These results have broader implications for public policy. They suggest, for example, that tariffs and source-based capital income taxes, which have been considered in the literature to be inappropriate in a small open economy, can, in fact, Pareto-improve welfare in a situation when non-linear income taxes are used for income redistribution.[4]

Recently, there have been several papers which focus on the relationship between incentive compatibility and the role of public policy. Boadway and Keen (1993), and Boadway and Marchand (1995) were the first to notice the importance of incentive compatibility in public policies. For example, Boadway and Marchand (1995) investigated the role of the public supply of semi-private goods under a non-linear income tax system. They found that supplying semi-public goods can relax the incentive constraint and make a non-linear income tax system more efficient.

Another line of literature has discussed the role of wage determination in a progressive tax system. Several authors found that if wages are endogenously determined, then the implications of a progressive tax system are quite different than if wages are exogenously determined (Feldstein 1973, Stern 1982, Stiglitz 1982). For example, Stiglitz (1982) analyzed an economy where two type of agents supply either skilled or unskilled labor, and these two types of labor are both used to produce one good. In his analysis, he found that if the government implements an optimal non-linear income tax, the optimal marginal tax rate on skilled workers can be negative.[5]

This paper is organized as follows. In Section 2, we set up the benchmark model. In Section 3, we will show that imposing a non-uniform commodity tax can Pareto-improve welfare even if the government is using a Pareto-efficient non-linear income tax system without a commodity tax. In Section 4, we will apply the model to the analysis of production efficiency in public production and show that introducing distortions in the public sector can Pareto-improve welfare. In Section 5, we will discuss the implications of our results for other public policies. Finally, in Section 6, we present

[4]See, e.g., Naito (1996).

[5]Since only one good is produced in Stiglitz model, the issue of optimal commodity taxes does not arise.

a short summary and suggested areas of future research.

2 Setup of the Model

2.1 Consumers

The basic model is an extension of Stiglitz (1982). There are two agent types, two factors, and two goods in this economy. The two agent types are skilled workers, denoted by superscript s, and unskilled workers, denoted by superscript u. Skilled workers supply skilled labor and unskilled workers supply unskilled labor. For simplicity, we assume that the population of both types of workers is the same and we normalize them to one.

We also assume that the utility function is weakly separable between consumption goods and labor supply, and that it is a strictly quasi-concave function with respect to (c_1^i, c_2^i, L^i):

$$V^i(U(c_1^i, c_2^i), L^i) \qquad\qquad (i = s, u),$$

where L^i is the labor supply of a worker of type i, and (c_1^i, c_2^i) is consumption of goods 1 and 2 by a worker of type i. We assume that good 1, good 2 and leisure are normal goods.[6] It is clear from the assumption of weak separability that the choice between c_1^i and c_2^i is independent of L^i if individual income is given. Thus, we can define the sub-indirect utility function $v(p + t, x)$ which measures utility given price and income:

$$\begin{aligned} & v(p+t, x) \equiv \max_{\{c_1^i, c_2^i\}} U(c_1^i, c_2^i) \\ & \text{subject to} \qquad c_1^i + (p+t)c_2^i \leq x, \end{aligned} \tag{1}$$

where p is the price of good 2, t is the commodity tax for good 2, and x is income. We assume that the price of good 1 is one, and that there is no commodity tax on good 1. From Roy's identity,

$$\frac{\partial v(p+t,x)}{\partial p} = -D(p+t,x)\frac{\partial v(p+t,x)}{\partial x}\ ,$$

[6] Normality assumption on goods and leisure is sufficient condition for so called a "single crossing property"and it plays an important role for determining the equilibrium. For a "single crossing property", see footnote 9.

where $D(p+t,x)$ is a conditional demand function of good 2 when the price is $p+t$ and income x is given. By using the sub-indirect utility function, we can define a conditional indirect utility function of type i agent as

$$V^i(v(p+t,x),L) \qquad (i=s,u) \ .$$

The conditional demand function is the maximized value from the consumption of two goods, when income x is given and labor supply is L.

2.2 Producers

There are two industries F_1 and F_2 in the private sectors of this economy. A production function of each industry is concave and exhibits constant returns to scale. We assume that each industry uses skilled labor and unskilled labor, and that each sector produces output y_1 and y_2, respectively. Thus,

$$y_1 = F^1(l_1^s, l_1^u) \ , \qquad\qquad y_2 = F^2(l_2^s, l_2^u), \tag{2}$$

where l_k^i is type i labor used in sector k $\{k=1,2\}$. Let w_s and w_u be wages for skilled workers and unskilled workers, respectively. Each industry maximizes its profit given the price of goods and wages. We assume that industry 2 is always skilled-labor intensive for any pair of $\{w_s, w_u\}$. Let $C_k(w_s,w_u)$ be a cost function to produce one unit of good k $\{k=1,2\}$. Since the economy is closed, we assume that both goods are produced in equilibrium without loss of generality. Then, perfect competition and constant returns to scale imply:

$$C_1(w_s,w_u) = 1 \ , \qquad\qquad C_2(w_s,w_u) = p \ . \tag{3}$$

By using Shepard's lemma, factor demands are

$$l_k^s = y_k \frac{\partial C_k(w_s,w_u)}{\partial w_s} \ , \qquad\qquad l_k^u = y_k \frac{\partial C_k(w_s,w_u)}{\partial w_u} \qquad (k=1,2) \ . \tag{4}$$

For a given price p , equations (3) determine w_s and w_u uniquely. Thus, we can write w_s, w_u and the ratio $w_u/w_s \equiv \Omega$ as a function of p :

$$w_s = w_s(p) \ , \quad w_u = w_u(p) \ , \quad \frac{w_u}{w_s} \equiv \Omega = \Omega(p) \ .$$

2.3 Government

The objective of the government is to achieve a Pareto-efficient allocation under the assumptions that the government cannot observe workers' types, but can observe each worker's total labor income. Therefore, in order to achieve a Pareto-efficient allocation, the government presents a menu of "tax contracts", so that individual workers self-select the contract which the government intended.[7] We define $T(\cdot)$ as a tax or subsidy function of the government and T_i as the amount of a tax or subsidy when the government observes the total income R^i. Then, T_i is:

$$T_i \equiv T(R^i) \ .$$

From the definition of T_i, the net income x^i of a worker i when she earns R^i becomes:

$$x^i = R^i - T_i \ .$$

When a type i worker earns R^j, her labor supply is $\frac{R^j}{w_i}$. The "tax contract" must satisfy incentive compatibility constraints. These constraints for workers are:[8]

$$V^s(v(p+t, x^s), \frac{R^s}{w_s}) \geq V^s(v(p+t, x^u), \frac{R^u}{w_s}) \tag{5}$$

$$V^u(v(p+t, x^u), \frac{R^u}{w_u}) \geq V^u(v(p+t, x^s), \frac{R^s}{w_u}) \ . \tag{6}$$

The first constraint means that a skilled worker has an incentive to work $\frac{R^s}{w_s} = L^s$, to report income R^s, and to receive a net income x^s instead of mimicking an unskilled worker, working $\frac{R^u}{w_s} = \frac{w_u L^u}{w_s}$, reporting income $R^u = w_u L^u$, and receiving a net income x^u. The second constraint is similarly for unskilled workers. [9] If the government provides an incentive-compatible menu $\{(R^u, x^u), (R^s, x^s)\}$, then all workers choose the "tax contract" which the government intended.

[7]Some might contend that we do not observe this type of tax system at all. However, "the Revelation Principle" proves that any tax system, or any mechanism, can be replicated by an incentive-compatible direct mechanism. Since real resource allocation is completely replicated by the incentive-compatible direct mechanism, we do not lose any generality by assuming this type of tax system. For an more detailed explanation about "the Revelation Principle", see Gibbons (1992) and Myerson (1985).

[8]We allow workers to pay negative tax. This means that the government pays a subsidy to workers.

[9]Furthermore, if we draw indifference curves of $V(v(p+t, x), \frac{R}{w_s})$ and $V(v(p+t, x), \frac{R}{w_u})$ with respect to (x, R), then the indifference curve of $V(v(p+t, x), \frac{R}{w_s})$ crosses the indifference curve of $V(v(p+t, x), \frac{R}{w_u})$ only once. In the literature of mechanism design, this is called a "single crossing property" .

The government budget constraint is:

$$T_s + T_u + t(D(p+t, x^s) + D(p+t, x^u)) \geq E(Q) \ . \tag{7}$$

The first two terms are the revenue and subsidies by a progressive income tax system, and the third term is the tax revenue from a commodity tax. $E(Q)$ is the cost of producing some amount of public goods.[10] By using the definition of x^s and x^u, we can rewrite the budget constraint in the following way:

$$w_s L^s + w_u L^u + t(D(p+t, x^s) + D(p+t, x^u)) \geq x^s + x^u + E(Q)$$

In order to examine the issue of production efficiency of public production, we assume that there is a public sector(public enterprise) which produces public goods in this economy.[11] The public enterprise produces some given amount of public goods using public production technology:

$$Q = G(l_q^s, l_q^u) \ , \tag{8}$$

where Q is the output of public goods, and l_q^s and l_q^u are the skilled and unskilled labor, respectively, used to produce those goods.[12] The government commands the public enterprise to produce some given amount of public goods Q by using the shadow prices z_u and z_s. The necessary money to produce public goods is transferred from the government.[13] Let $C_q(z_s, z_u, Q)$ be the cost function for the public enterprise. Then, the labor demand of the public enterprise is:

$$l_q^s = \frac{\partial C_q(z_s, z_u, Q)}{\partial z_s} \ , \qquad l_q^u = \frac{\partial C_q(z_s, z_u, Q)}{\partial z_u} \ . \tag{9}$$

[10] In this paper, without loss of generality, we assume that the level of public goods is exogenous and fixed. Although it is possible to endogenize the level of public goods in this model, our conclusion does not change.

[11] The existence of a public sector is not essential for the result of optimality of non-uniform commodity taxation. We assume the existence of a public sector in order to analyze the issue of production efficiency of public production.

[12] We assume that the isoquant curve of the production function is strictly convex, so that the cost minimizing factor-demand is uniquely determined.

[13] We assume that there is no asymmetric information between the public sector and the government to focus on the issue of production efficiency. It is standard to adopt this assumption when the issue is production efficiency of public production.

The expenditure to produce the public goods is:

$$E(Q) \equiv l_q^s w_s + l_q^u w_u \ .$$

Since the factor demand function is homogeneous of degree zero with respect to z_s, z_u, and since the cost of producing public goods is financed by the government, what is important for the government is the shadow price ratio for the public enterprise. Thus, let $\frac{z_u}{z_s} = (1+\zeta)\frac{w_u}{w_s} = (1+\zeta)\Omega$ be the shadow price ratio by which the government tells public enterprises to minimize their costs. Then, l_q^s and l_q^u become a function of $(1+\zeta)\Omega$. In this case, ζ denotes the degree of distortion in public production. If ζ is equal to zero, public production is efficient, i.e., the marginal rate of transformation between skilled and unskilled labor in the public sector is equal to that of the private sectors. If ζ is not equal to zero, there is a distortion in public production.

2.4 Equilibrium

As for labor markets, we assume that labor is perfectly mobile among the two private sectors and one public sector. Thus, the labor market equilibrium conditions are:

$$L^s = l_q^s + l_1^s + l_2^s \ , \qquad L^u = l_q^s + l_1^u + l_2^u \ . \tag{10}$$

The goods market equilibrium conditions are:

$$y_1 = c_1^u + c_1^s \ , \qquad y_2 = c_2^u + c_2^s \equiv D(p+t, x^s) + D(p+t, x^u) \ . \tag{11}$$

By Walras' law, we can ignore the equilibrium condition of good 1 .

For the analysis in Section 3, it is useful to know how $\{p, w_s, w_u\}$ are determined in the economy. For this purpose, we use the production-possibility frontier of the private sector, conditional on a given total labor force in the private sector. Because technology is convex and factor intensity is different between the two sectors, the production possibility set is a strictly convex set. Thus, once the price ratio between good 1 and good 2 and the total labor force in the private sector are given, the amount of goods 1 and 2 produced is uniquely determined. Let $\bar{l^s} \equiv l_1^s + l_2^s$ and $\bar{l^u} \equiv l_1^u + l_2^u$ be the total labor force of skilled and unskilled labor, respectively, in the private sector. Then, we can write the output function of good 2 as:

$$y_2 = Y(p; \bar{l^s}, \bar{l^u}) \ .$$

Since the production possibility set is strictly convex,

$$\frac{\partial Y(p; \bar{l^s}, \bar{l^u})}{\partial p} > 0$$

if $\quad Y > 0$.

From the Rybczynski theorem, if the skilled(unskilled) labor force in the private sector increases, then the production of the skilled-labor(unskilled-labor) intensive good will increase, and the production of the unskilled-labor(skilled-labor) intensive good will decrease given a fixed price, p. Thus:

$$\frac{\partial Y(p; \bar{l^s}, \bar{l^u})}{\partial \bar{l_s}} > 0, \qquad \frac{\partial Y(p; \bar{l^s}, \bar{l^u})}{\partial \bar{l^u}} < 0 \ .$$

The market equilibrium price p is determined from the following equation:

$$D(p + t, x^s) + D(p + t, x^u) = Y(p; \bar{l^s}, \bar{l^s}) \ . \tag{12}$$

Once price p is determined, then wages are determined from equation (3). From the Stolper and Samuelson (1941) theorem, the effect of the change of price on wages is:[14]

$$w_s'(p) > 0, \quad w_u'(p) < 0, \quad \Omega'(p) < 0.$$

Using the equilibrium conditions of the labor market, we can rewrite the market equilibrium condition (12) as:

$$D(p + t, x^s) + D(p + t, x^u) = Y(p\,; L^s - l_q^s, L^u - l_q^s) \ . \tag{13}$$

Since l_q^s and l_q^u are functions of $(1 + \zeta)\Omega(p)$, the equilibrium price p is a function of L^s, L^u, t, x^s, x^u, t and ζ. We denote the equilibrium price function as $p(x^s, x^u, L^s, L^u, t, \zeta)$.

[14] Intuitively, if the producer price of good 2 increases, then the output of good 2 will increase and the output of good 1 will decrease. Since sector 1 is unskilled-labor intensive and sector 2 is skilled-labor intensive, the wages of skilled labor must increase and the wages of unskilled labor must decrease to restore equilibrium in the labor market.

3 Optimality of A Non-Uniform Commodity tax

This section examines whether additionally imposing a commodity tax can Pareto-improve welfare when the government is using a non-linear Pareto-efficient tax system without a commodity tax. For this purpose, we first characterize the Pareto-efficient non-linear labor income tax system for a given level of commodity tax and a given distortion in the public sector. Then, we check whether marginally increasing a commodity tax from zero can Pareto-improve welfare when the non-linear labor income tax is optimally adjusted.

The Pareto-efficient non-linear income tax system for a given level of commodity tax and a given distortion in the public sector can be obtained by solving the following programming problem:

Main program

$$\max_{\{L^s, L^u, x^s, x^u\}} V^s(v(p(\cdot)+t, x^s), L^s)$$

subject to:

$$V^u(v(p(\cdot)+t, x^u), L^u) \geq \bar{U^u} \ , \tag{MUC}$$

$$V^s(v(p(\cdot)+t, x^s), L^s) \geq V^s(v(p(\cdot)+t, x^u), \Omega(\cdot)L^u) \ , \tag{ICS}$$

$$V^u(v(p(\cdot)+t, x^u), L^u) \geq V^u(v(p(\cdot)+t, x^s), \frac{L^s}{\Omega(\cdot)}) \tag{ICU}$$

$$\begin{aligned} & w_s(p(\cdot))L^s + w_u(p(\cdot))L^u + tD(p(\cdot)+t, x^s) + tD(p(\cdot)+t, x^u) \\ & \quad \geq x^s + x^u + w_s l^s_q((1+\zeta)\Omega(\cdot)) + w_u l^u_q((1+\zeta)\Omega(\cdot)) \ , \end{aligned} \tag{BC}$$

where , $\quad p(\cdot) = p(L^s, L^u, x^s, x^u, t, \zeta) \quad$ and $\quad \Omega(\cdot) = \Omega(p(\cdot))$

t and ζ are given.

In the above programming problem, MUC is the utility constraint for unskilled workers. The Pareto-efficient allocation can be obtained by maximizing the utility of one type of workers subject to the constraint that the utility of the other type of workers is above some level $\bar{U^u}$.[15] ICS is the incentive compatibility constraint for skilled workers and ICU is the incentive compatibility constraint

[15] Since controlling R is equivalent to controlling L, we use x and L as control variables.

for unskilled workers. BC is the government budget constraint. The government chooses L^s, L^u, x^s, and x^u, so that the Pareto-efficient allocation is achieved subject to MUC, ICU, ICS, and BC. In this optimization problem, the government knows that p is a function of $\{L^s, L^u, x^s, x^u, t, \zeta\}$, and that $\{w_s, w_u, \Omega\}$ are functions of p.

Let γ, λ_1, λ_2, and λ_3 be the Lagrangian multipliers of MUC, ICS, ICU, and BC, respectively and $\mathcal{L}(L^s, L^u, x^s, x^u, \mu, \lambda_1, \lambda_2, \lambda_3; t, \zeta)$ be the Lagrangian function. For simplicity, let us initially assume that there is no public enterprise in this economy and that the government collects taxes for only income redistribution.[16] Let $\{L^s(t), L^u(t), x^s(t), x^u(t)\}$ be the solution of the *Main program*, and $W(t)$ be the maximized value of the objective function of the *Main program*:

$$
\begin{aligned}
&W(t) \equiv V^s(v(p(\cdot)+t, x^s(t)), L^s(t)) \\
&\text{where } \; p(\cdot) = p(x^s(t), x^u(t), L^s(t), L^u(t), t).
\end{aligned}
$$

Then, we can check whether imposing a commodity tax can Pareto-improve welfare in this economy by calculating $\frac{dW}{dt}\big|_{t=0}$. For convenience, we define:

$$
\begin{aligned}
&\frac{\partial V^i(v(p+t, x^i), L^i)}{\partial\{v(p+t, x^i)\}} \equiv V_1^{ii}, && \frac{\partial V^i(v(p+t, x^i), L^i)}{\partial L^i} \equiv V_2^{ii} \\
&\frac{\partial V^i(v(p+t, x^j), \frac{w_j L^j}{w_i})}{\partial\{v(p+t, x^j)\}} \equiv V_1^{ij} \quad (i \neq j), && \frac{\partial V^i(v(p+t, x^j), \frac{w_j L^j}{w_i})}{\partial\{\frac{w_j L^j}{w_i}\}} \equiv V_2^{ij} \quad (i \neq j) \\
&\frac{\partial v(p+t, x^i)}{\partial(p+t)} \equiv v_p(p+t, x^i), && \frac{\partial v(p+t, x^i)}{\partial x^i} \equiv v_x(p+t, x^i) \\
&\text{i, j=s, u} \;\; .
\end{aligned}
$$

$$
Y_p = \frac{\partial Y(p; \bar{l^s}, \bar{l^u})}{\partial p} \qquad Y_{\bar{l^i}} = \frac{\partial Y(p; \bar{l^s}, \bar{l^u})}{\partial \bar{l^i}}
$$

$$
D_p(p+t, x^i) = \frac{\partial D(p+t, x^i)}{\partial(p+t)} \qquad D_x(p+t, x^i) = \frac{\partial D((p+t, x^i)}{\partial x^i} \; .
$$

[16] In Section 4, we will analyze the case where there is a public enterprise in the economy and the government chooses optimal ζ.

From the envelope theorem, $\frac{dW}{dt}\Big|_{t=0}$ is equal to :

$$\begin{aligned}
\frac{dW}{dt}\bigg|_{t=0} &= \frac{\partial \mathcal{L}}{\partial t}\bigg|_{t=0} \\
&= V_1^{ss} v_p(p, x^s) + \gamma V_1^{uu} v_p(p, x^u) \\
&\quad + \lambda_1 \{V_1^{ss} v_p(p, x^s) - V_1^{su} v_p(p, x^u)\} + \lambda_2 \{V_1^{uu} v_p(p, x^u) - V_1^{us} v_p(p, x^s)\} \\
&\quad + \lambda_3 \{D(p, x^s) + D(p, x^u)\} + \frac{\partial p}{\partial t}\Delta \ ,
\end{aligned}$$

$$\begin{aligned}
\text{where} \quad \Delta &= V_1^{ss} v_p(p+t, x^s) + \gamma V_1^{uu} v_p(p+t, x^l) \\
&\quad + \lambda_1 \{V_1^{ss} v_p(p+t, x^s) - V_1^{su} v_p(p+t, x^u) - V_2^{su} \Omega'(p) * L^u\} \\
&\quad + \lambda_2 \{V_1^{uu} v_p(p+t, x^u) - V_1^{us} v_p(p+t, x^s) - V_2^{us} L^s \frac{-\Omega'(p)}{\Omega^2}\} \\
&\quad + \lambda_3 \{w_s'(p) L^s + w_u'(p) L^u - w_s'(p) l_q^s - w_u'(p) l_q^u - w_s \frac{\partial l_q^s}{\partial p} - w_u \frac{\partial l_q^u}{\partial p}\} \ .
\end{aligned}$$

Before we examine this condition, it would be useful to consider the case in which the marginal cost of producing goods is constant and p is fixed. (This is a case of a small open economy with no tariff.) By using the first order conditions for x^s and x^u, $\frac{\partial p}{\partial t} = 0$, $\frac{\partial p}{\partial x^s} = 0$ and $\frac{\partial p}{\partial x^u} = 0$, equation (**??**) becomes

$$\frac{dW}{dt}\bigg|_{t=0} = 0 \ . \tag{14}$$

This equation implies that, at the margin, the government does not have a motivation to impose a non-uniform commodity tax. Thus, we obtain the result of Atkinson and Stiglitz (1976) in our framework.

Result 1 (Atkinson and Stiglitz 1976).

If the government can implement a non-linear labor income tax and the utility functions are weakly separable between leisure and goods, then uniform commodity taxes are optimal when the prices of goods are constant.

How will this result change if we assume that p is not fixed? What we find is that uniform commodity taxation is no longer optimal when p is variable. In particular,

Theorem 1 (Pareto-improving non-uniform commodity taxation).
Consider an economy where the government is using a Pareto-efficient non-linear labor income tax system without a commodity tax, and where the utility function of the workers is weakly separable between goods and leisure. If the non-liner labor income tax system is "redistributive", then imposing a commodity tax on skilled labor-intensive goods Pareto-improves welfare. If the non-linear income tax is "regressive", then imposing a commodity tax on a unskilled labor intensive goods Pareto-improves welfare.

If a non-linear income tax is "first best", then there is no welfare gain from imposing a commodity tax.

The proof of the theorem goes as follows. From equation (**??**), the effect of a commodity tax on welfare is (see Appendix 1):

$$\left.\frac{dW}{dt}\right|_{t=0} = \Psi_1 \{\lambda_1 V_2^{su}\Omega'(p)L^u - \lambda_2 V_2^{us} L^s \frac{\Omega'(p)}{\Omega^2}\} \tag{15}$$

$$\text{where} \quad \Psi_1 = \frac{h_p(p, v^s) + h_p(p, v^u)}{h_p(p, v^s) + h_p(p, v^u) - Y_p} \ ,$$

and $h_p(p, v^i)$ is the partial derivative of the Hicksian demand function with respective to its own price.

The above equation (15) has a clear economic interpretation. First, the term Ψ_1 measures how the producer price will change when the government imposes a commodity tax on good 2 and compensates income so that the utilities of workers are kept constant.[17]

[17] If the government compensate income so that the utility of workers are kept constant when the government impose a commodity tax, then the following equation will hold:

$$h(p + t, v^s) + h(p + t, v^u) = Y(p, L^s, L^u)$$

By totally differentiating the above equation with respect to p and t, and solving for $\frac{dp}{dt}$, we obtain the formula of Ψ_1.

Second, the terms λ_1 and λ_2 show how much welfare is increased when the incentive compatibility constraint is relaxed. Third, the terms $V_2^{su}\Omega'(p)L^u$ and $V_2^{us}L^s\frac{\Omega'(p)}{\Omega^2}$ show how much the incentive compatibility constraint is relaxed by a change of wage ratio through a change of the producer price. Therefore, equation (15) show how much welfare is increased through the relaxation of the incentive compatibility constraint by imposition of a commodity tax on good 2.

Since from our assumption the utility function is strictly quasi-concave, the partial derivative of the Hicksian demand function with respect to its own price is strictly negative. Y_p is positive from the shape of the production possibility frontier. Thus, Ψ_1 is strictly positive. This implies that the effect of a commodity tax on welfare depends on the values of λ_1 and λ_2, while the values of λ_1 and λ_2 depend on the structure of the non-linear income tax system at $t = 0$. There are three cases of interest for λ_1 and λ_2:[18][19]

1. Redistributive case: $\lambda_1 > 0 \quad \lambda_2 = 0$.

   This is the case which most of literature has focused on, and which Stiglitz (1982) called the 'normal case'. In this case, ICS is binding and ICU is not binding. This is because the government sets the level of utility of unskilled workers $\bar{U}$ relatively high and needs to redistribute income from skilled workers to unskilled workers. As a result, skilled workers want to mimic unskilled workers but unskilled workers do not want to mimic skilled workers. In this situation, the effect of imposing a commodity tax on welfare is

$$\left.\frac{dW}{dt}\right|_{t=0,} = \Psi_1\lambda_1 V_2^{su}\Omega'(p)L^u > 0 \ .$$

   Therefore, imposing a commodity tax on a skilled labor-intensive good Pareto-improves welfare.

2. Regressive case: $\lambda_1 = 0 \quad \lambda_2 > 0$.

   In this case, the situation is converse and imposing a commodity tax on unskilled-labor intensive good Pareto-improves welfare.

[18]Without loss of generality we assume that the equilibrium wage of skilled labor is greater than the wage of unskilled labor at $t = 0$.

[19]There is a forth case so called "bunching" where both ICS and ICU are binding. In this case, the single crossing property of the utility function implies that both types of workers earn the same gross income and receive the same net income and the effect of non-uniform commodity taxation is ambiguous.

3. First best case: $\lambda_1 = 0 \quad \lambda_2 = 0$.

   In this case, the allocation is the first-best and, thus, imposing a commodity tax does not change welfare.

As we argued in the Introduction, this result is quite different from results indicated in the previous literature (Atkinson and Stiglitz 1976, 1980). Our result suggests that if a non-linear income tax distorts labor and consumption decisions, then it is possible to Pareto-improve welfare by distorting the market of consumption goods.

The economic intuition of our result is as follows. Consider a situation where the government is using a Pareto-efficient "redistributive" non-linear income tax system but does not impose any commodity taxes. If the government begins to impose a commodity tax on a skilled labor-intensive good, then the wage of unskilled labor will increase and the wage of skilled labor will decrease. Since the wage differential between the skilled and the unskilled becomes smaller, the government can let the unskilled work more and earn more keeping the utility of the unskilled constant, because such an allocation is no longer mimicked by the skilled due to the decrease of the wage differential and the relaxation of the incentive compatibility constraint. Since the unskilled earns more, the tax burden of the skilled can be reduced. This reduction of the tax burden has a first order effect on welfare, whereas the distortion of a commodity tax is of second order. Therefore, imposing a commodity tax can Pareto-improve welfare.

An interesting feature of our result is that the more elastic the demand function, the larger the welfare gain. This is opposite to results reported in the previous literature. Generally, we are accustomed to thinking that if the elasticity is high, then the welfare loss is high. However, in this case if the elasticity is high, then the welfare gain is high because it greatly relaxes the incentive compatibility constraint.

4 Production Efficiency of Public Production

In this section, we apply the insights of the previous section to the issue of production efficiency in public enterprise. Consider a situation where the government solved the *Main program* under the assumption of efficient public production ($\zeta = 0$) and found that imposing a commodity tax is Pareto-improving. Suppose also that the government chose a commodity tax t at optimal levels under efficient public production. Then, our next question is whether or not introducing a distortion in the public sector can Pareto-improve welfare.

Let $\{L^s(\zeta), x^s(\zeta), L^u(\zeta), x^u(\zeta), t(\zeta)\}$ be the solution of the first order conditions for x^s, x^u, L^s, L^u, t and the constraints of MUC, ICS and BC for a given ζ. Then the maximized value of the *Main program* for a given ζ becomes:

$$W(\zeta) \equiv V^s(v(p(\cdot) + t(\zeta), x^s(\zeta)), L^s(\zeta))$$
$$\text{where} \quad p(\cdot) = p(x^s(\zeta), x^u(\zeta), L^s(\zeta), L^u(\zeta), t(\zeta), \zeta) \ .$$

Then, by calculating $\left.\frac{dW(\zeta)}{d\zeta}\right|_{\zeta=0}$, we can check whether or not introducing a distortion in the public sector can Pareto-improve welfare. Using the same technique used in the previous section, we obtain the following lemma.

Lemma 1.

Consider an economy where the government is using a Pareto-efficient non-linear income tax and a commodity tax for income redistribution under efficient public production. If the government introduces a distortion to public production, then the effect on welfare is:

$$\left.\frac{dW(\zeta)}{d\zeta}\right|_{\zeta=0} = \Psi_2\{-\lambda_1 V_2^{su}\Omega' L^u + \lambda_2 V_2^{us}\frac{\Omega'}{\Omega^2}L^s\} \ , \tag{16}$$
$$\textit{where} \quad \Psi_2 = \frac{-Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial \zeta} - Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial \zeta}}{-Y_p + Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial p} + Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial p}} \ .$$

Proof: See Appendix 2 .

In the above lemma, $\frac{\partial l_q^s}{\partial p} \equiv \frac{\partial l_q^s}{\partial \Omega}\frac{\partial \Omega}{\partial p} < 0$, $\frac{\partial l_q^u}{\partial p} \equiv \frac{\partial l_q^u}{\partial \Omega}\frac{\partial \Omega}{\partial p} > 0$, $\frac{\partial l_q^s}{\partial \zeta} > 0$, and $\frac{\partial l_q^u}{\partial \zeta} < 0$ from the assumption on the shape of isoquant. From the Rybczynski theorem, $Y_{\bar{l^s}} > 0$ and $Y_{\bar{l^u}} < 0$. Thus, Ψ_2 is strictly positive. Therefore, the effect on welfare of a distortion in the public sector depends on the values of λ_1 and λ_2. Based on the same argument in Section 3, we obtain the following theorem:

Theorem 2 (Distortion in public production for income redistribution).
Consider an economy where the government is using a Pareto-efficient non-linear labor income tax system with a commodity tax under the efficient public production. If the non-liner labor income tax system is "redistributive(regressive)", then decreasing(increasing) the shadow-wage ratio between unskilled and skilled labor, $\frac{z_u}{z_s}$*, from the level of the market wage ratio Pareto-improves welfare.*

If a non-linear income tax is 'first best', then there is no welfare gain from distortion in the public sector.

This result differs from results previously reported in the literature about the production efficiency of public enterprises (Diamond and Mirrlees 1971). Theorem 2 implies that production efficiency of public enterprise is not necessarily desirable.

The economic intuition of the above theorem is as follows. Consider the situation where a Pareto-efficient non-linear income tax system under efficient public production is redistributive. Suppose that the government decreases the shadow-wage ratio between unskilled and skilled labor for public enterprises, inducing them to employ more unskilled workers and few skilled workers than the efficient production level. If public enterprises employ more unskilled workers, the unskilled labor force in the private sectors will decrease and the skilled labor force will increase. From the Rybczynski theorem, this implies that the output of unskilled labor-intensive goods will decrease and the output of skilled labor-intensive goods will increase. Then, in goods markets the relative price of skilled labor intensive-goods will decrease. From the Stolper-Samuelson theorem, we know that an increase of the relative price of unskilled labor-intensive goods increases the wage ratio between unskilled and skilled labor. Finally, such an increase of the relative wages between unskilled and skilled labor relaxes the incentive compatibility constraint and Pareto-improves welfare for the same reasons as shown in the previous section.

Another interpretation of our proposition is that by hiring more unskilled workers in the public sector, the government can create a scarcity of unskilled worker in the private sector and indirectly redistribute income from skilled to unskilled workers. By this indirect income redistribution, the government can mitigate the incentive problem of redistributive income taxation.[20]

This theorem has an important practical implication for the governments of developing countries. In many of these developing countries, the government is heavily engaged in public production. The previous literature in public finance (Diamond and Mirrlees 1971) says that public production should minimize cost based on market prices for all factors as long as the government has enough instruments to control all factor prices. However, our theorem shows when a government needs to use non-linear income taxation due to asymmetric information, then redistribution through distorted factor inputs in public production becomes attractive.

5 Policy Implication

Theorems 1 and 2 have important implications for public policy. They suggest that many seemingly peculiar policies can be explained as mechanisms to relax the incentive problem of income redistribution. For example, Naito (1996) showed that governments can Pareto-improve welfare either by imposing a tariff or providing a production subsidy on unskilled labor-intensive goods in a small open economy when the government is using a redistributive non-linear income tax system. This is because tariffs or a production subsidies can change the wage ratio and reduce the incentive problem of income redistribution. This result is in sharp contrast to that of the Stolper-Samuelson theorem in

[20]One might wonder why we obtained different results from those in the previous literature. Diamond and Mirrlees (1971) showed that if the government has instruments to control the price of all goods (including factors) freely, then production efficiency of public production is optimal. However, in our framework, the government cannot control factor prices freely because of asymmetric information and incentive compatibility problems. Thus, the assumption of Diamond and Mirrlees does not hold in our framework.

international trade theory. That theorem says that imposing a tariff necessarily benefits one group and hurts the other group. However, our result shows that imposing a tariff *Pareto-improves* welfare if a country is using a non-linear income tax system to redistribute income. Since most countries use a progressive income tax system to some extent, our result suggests that governments have a strong motivation to impose tariffs, even in a small open economy.

A second implication of our result is that source-based capital income taxation may be appropriate even in a small open economy. In the public finance literature, it is argued that source-based capital income taxation is inefficient in a small open economy (Gordon 1986; Razin and Sadka 1991). For example, Gordon (1986) argued that because the tax burden on the mobile factor (capital) in a small open economy is shifted completely to the immobile factor (labor), it is more efficient to tax labor directly since the tax burden is shifted to labor anyway.

Although the above argument is true when an economy consists of one representative agent, the result of the present research suggests that the above argument does not hold when there are several agents in the economy and the government is concerned with efficient income redistribution under asymmetric information between the government and agents. For example, if there are two types of workers in the economy, skilled and unskilled, and if the government is concerned with Pareto-efficient income redistribution, then the government will use a source-based capital income tax even in the small open economy, because the source-based capital income tax can change the wages of workers through the general equilibrium effect of factor markets, thus shifting the tax-burden from one type of workers to the other. This shift of the tax burden through a source-based capital income tax reduces the disincentive effect of a redistributive labor income tax and makes the whole economy more efficient.

6 Conclusion

In this paper, we re-examined two well-established results (Atkinson and Stiglitz 1976 and Diamond and Mirrlees 1971) in the public finance literature by introducing the *production side* in the analysis and considering the use of nonlinear income taxes rather than separate factor taxes on each skill level due to the government's inability to observe skill types. In a general equilibrium model with two agent types, two factors, and two goods, we first showed that the government can Pareto-improve welfare by distorting the relative consumption prices. Second, we proved that introducing distortions in public production also can Pareto-improve welfare.

The key insight of this paper is that these distortions help relax the incentive problem in income redistribution which arises under asymmetric information. This insight sheds new light on many public policies, such as tariffs and source-based capital income taxation in an open economy.

A Appendix

A.1 Derivation of equation (15)

From Roy's identity, equation (**??**) becomes:

$$\begin{aligned}\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} &= -D(p,x^s)\{V_1^{ss}v_x(p,x^s) + \lambda_1 V_1^{ss}v_x(p,x^s) - \lambda_2 V_1^{us}v_x(p,x^s) - \lambda_3\} \\ &\quad - D(p,x^u)\{\gamma V_1^{uu}v_x(p,x^u) - \lambda_1 V_1^{uu}v_x(p,x^u) + \lambda_2 V_1^{uu}v_x(p,x^u) - \lambda_3\} + \frac{\partial p}{\partial t}\Delta\ .\end{aligned}$$

From the first order conditions for x^s and x^u, the above equation is equal to:

$$\begin{aligned}\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} &= \Delta D(p,x^s)\frac{\partial p}{\partial x^s} + \Delta D(p,x^u)\frac{\partial p}{\partial x^u} + \Delta\frac{\partial p}{\partial t} \\ &= \Delta\{D(p,x^s)\frac{\partial p}{\partial x^s} + D(p,x^u)\frac{\partial p}{\partial x^u} + \frac{\partial p}{\partial t}\}\ .\end{aligned}$$

By calculating $\frac{\partial p}{\partial x^s}$, $\frac{\partial p}{\partial x^u}$ and $\frac{\partial p}{\partial t}$ from equation (13), the above equation becomes:

$$\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} = -\Delta\left\{\frac{D_x(p,x^s)D(p,x^s) + D_x(p,x^u)D(p,x^u) + D_p(p,x^s) + D_p(p,x^u)}{D_p(p,x^s) + D_p(p,x^u) - Y_p}\ .\right\}$$

By using the Slutsky equation, we obtain:

$$\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} = -\Delta\left\{\frac{h_p(p,v(p,x^s)) + h_p(p,v(p,x^u))}{D_p(p,x^s) + D_p(p,x^u) - Y_p}\ .\right\} \tag{17}$$

Next, we need to calculate Δ. From the definition of Δ and Roy's identity again, Δ is equal to:

$$\begin{aligned}\Delta = &-D(p,x^s)\{V_1^{ss}v_x(p,x^s)+\lambda_1 V_1^{ss}v_x(p,x^s)-\lambda_2 V_1^{us}v_x(p,x^s)\}\\ &-D(p,x^u)\{\gamma V_1^{uu}v_x(p,x^u)-\lambda_1 V_1^{su}v_x(p,x^u)+\lambda_2 V_1^{uu}v_x(p,x^u)\}\\ &+\lambda_3\{w_s'(p)L^s+w_u'(p)L^u\}\\ &-\lambda_1 V_2^{su}\Omega'(p)L^u+\lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ .\end{aligned}$$

Using the first order conditions for x^s and x^u again, equation (**??**) becomes:

$$\begin{aligned}\Delta = &-D(p,x^s)\{\lambda_3-\frac{\partial p}{\partial x^s}\Delta\}-D(p,x^u)\{\lambda_3-\frac{\partial p}{\partial x^u}\Delta\}\\ &+\lambda_3\{w_s'(p)L^s+w_u'(p)L^u\}\\ &-\lambda_1 V_2^{su}\Omega'(p)L^u+\lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ .\end{aligned}$$

Then, we need to calculate:

$$w_s'(p)L^s+w_u'(p)L^u \ .$$

We note that perfect competition implies

$$w_s(p)L^s+w_u(p)L^u=y_1+py_2$$

Thus the magnitude of $w_s'(p)L^s+w_u'(p)L^u$ is equivalent to the change of total revenue when the labor force is fixed. If labor force is fixed, the shape of the production possibility frontier is uniquely determined. Thus, knowing the output of y_1 and y_2 is equivalent to solving the following problem:

$$\begin{aligned}&\pi(p)\equiv\max_{\{y_1,y_2\}} y_1+py_2\\ &\text{subject to}\qquad (y_1,y_2)\in\bar{Y}(L^s,L^u)\ ,\end{aligned}$$

where $\bar{Y}(L^s,L^u)$ is a production possibility set. Therefore from the envelope theorem,

$$w_s'(p)L^s+w_u'(p)L^u=\frac{d\pi(p)}{dp}=y_2.$$

Since $y_2=D(p,x^s)+D(p,x^u)$ in the equilibrium,

$$w_s'(p)L^s+w_u'(p)L^u=D(p,x^s)+D(p,x^u)\ . \tag{18}$$

Thus, by using equation (18), equation (**??**) becomes:

$$\{1 - D(p, x^s)\frac{\partial p}{\partial x^s} - D(p, x^u)\frac{\partial p}{\partial x^u}\}\Delta = -\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ .$$

From the definition of $\frac{\partial p}{\partial x^s}$ and $\frac{\partial p}{\partial x^u}$, the above equation becomes

$$\left\{\frac{D_p(p, x^s) + D_p(p, x^u) - Y_p + D_x(p, x^s)D(p, x^s) + D_x(p, x^u)D(p, x^u)}{D_p(p, x^s) + D_p(p, x^u) - Y_p}\right\}\Delta$$
$$= -\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ .$$

Using the Slutsky equation again, the above equation becomes:

$$\left\{\frac{h_p(p, v(p, x^s)) + h_p(p, v(p, x^u)) - Y_p}{D_p(p, x^s) + D_p(p, x^u) - Y_p}\right\}\Delta$$
$$= -\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ .$$

Solving the above equation for Δ implies that

$$\begin{aligned}\Delta = &\frac{D_p(p, x^s) + D_p(p, x^u) - Y_p}{h_p(p, v(p, x^s)) + h_p(p, v(p, , x^u)) - Y_p} \\ &\times \{-\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s\} \ .\end{aligned} \tag{19}$$

Finally, by substituting equation (19) for Δ, equation (17) becomes

$$\begin{aligned}\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} = &\frac{h_p(p, v(p, x^s)) + h_p(p, v(p, x^u))}{h_p(p, v(p, x^s)) + h_p(p, v(p, , x^u)) - Y_p} \\ &\times \{-\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s\} \ .\end{aligned} \tag{20}$$

This is equation (15).

A.2 The proof of lemma 1

From the first order condition of cost minimization in public sector, $w_s\frac{\partial l_q^s}{\partial \zeta} + w_u\frac{\partial l_q^s}{\partial \zeta}$ is zero at $\zeta = 0$. Thus,

$$\left.\frac{dW(\zeta)}{d\zeta}\right|_{\zeta=0} = \left.\frac{\partial \mathcal{L}}{\partial \zeta}\right|_{\zeta=0} = \frac{\partial p}{\partial \zeta}\Delta \ . \tag{21}$$

Next, we need to calculate Δ. By the definition of Δ,

$$\begin{aligned}
\Delta = & V_1^{ss} v_p(p+t, x^s) + \gamma V_1^{uu} v_p(p+t, x^u) \\
& + \lambda_1 \{ V_1^{ss} v_p(p+t, x^s) - V_1^{su} v_p(p+t, x^u) - V_2^{su} \Omega'(p) * L^u \} \\
& + \lambda_2 \{ V_1^{uu} v_p(p+t, x^u) - V_1^{us} v_p(p+t, x^s) - V_2^{us} L^s \frac{-\Omega'(p)}{\Omega^2} \} \\
& + \lambda_3 \{ w_s'(p) L^s + w_u'(p) L^u + t D_p(p+t, x^s) + t D_p(p+t, x^u) \\
& - w_s'(p) l_q^s - w_u'(p) l_q^u - w_s \frac{\partial l_q^s}{\partial p} - w_u \frac{\partial l_q^u}{\partial p} \} \ .
\end{aligned}$$

Using the technique for equation (15), we obtain the following equation for Δ:

$$\begin{aligned}
\Delta = & -D(p, x^s) \{ V_1^{ss} v_x(p+t, x^s) + \lambda_1 V_1^{ss} v_x(p+t, x^s) - \lambda_2 V_1^{us} v_x(p, x^s) \} \\
& - D(p, x^u) \{ \gamma V_1^{uu} v_x(p+t, x^s) - \lambda_1 V_1^{su} v_x(p+t, x^u) + \lambda_2 V_1^{uu} v_x(p+t, x^u) \} \\
& + \lambda_3 \{ w_s'(p) L^s + w_u'(p) L^u - w_s'(p) l_q^s - w_u'(p) l_q^u - w_s \frac{\partial l_q^s}{\partial p} - w_u \frac{\partial l_q^u}{\partial p} \} \\
& \lambda_3 t D_p(p+t, x^s) + \lambda_3 t D_p(p+t, x^u) \\
& - \lambda_1 V_2^{su} \Omega'(p) L^u + \lambda_2 V_2^{us} \frac{\Omega'(p)}{\Omega^2} L^s.
\end{aligned}$$

As for $w_s'(p) L^s + w_u'(p) L^u - w_s'(p) l_q^s - w_u'(p) l_q^u - w_s \frac{\partial l_q^s}{\partial p} - w_u \frac{\partial l_q^u}{\partial p}$, by using equation (18) and the fact that $w_s \frac{\partial l_q^s}{\partial p} + w_u \frac{\partial l_q^u}{\partial p} = 0$, we obtain:

$$w_s'(p) L^s + w_u'(p) L^u - w_s'(p) l_q^s - w_u'(p) l_q^u - w_s \frac{\partial l_q^s}{\partial p} - w_u \frac{\partial l_q^u}{\partial p} = D(p+t, x^s) + D(p+t, x^u) \ .$$

Combined with the first order condition for x^s and x^u, Δ becomes

$$\begin{aligned}
\Delta = & \lambda_3 t D(p+t, x^s) D_x(p+t, x^s) + \lambda_3 D(p+t, x^s) \frac{\partial p}{\partial x^s} \Delta \\
& + \lambda_3 t D(p+t, x^u) D_x(p+t, x^u) + D(p+t, x^u) \frac{\partial p}{\partial x^u} \Delta \\
& + \lambda_3 t D_p(p+t, x^s) + \lambda_3 t D_p(p+t, x^u) \\
& - \lambda_1 V_2^{su} \Omega'(p) L^u + \lambda_2 V_2^{us} \frac{\Omega'(p)}{\Omega^2} L^s \ .
\end{aligned}$$

On the other hand, from the first order condition for t,

$$\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{\zeta=0} = \lambda_3 t D(p+t,x^s)D_x(p+t,x^s) + \lambda_3 D(p+t,x^s)\frac{\partial p}{\partial x^s}\Delta$$
$$+ \lambda_3 t D(p+t,x^u)D_x(p+t,x^u) + D(p+t,x^u)\frac{\partial p}{\partial x^u}\Delta$$
$$+ \lambda_3 t D_p(p+t,x^s) + \lambda_3 t D_p(p+t,x^u) + \frac{\partial p}{\partial t}\Delta = 0 \ .$$

Therefore, equation (**??**) becomes

$$\Delta = -\frac{\partial p}{\partial t}\Delta - \lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s \ . \tag{22}$$

Solving for Δ,

$$\begin{aligned}\Delta &= \frac{1}{1+\frac{\partial p}{\partial t}}\{-\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2 s^{us}\frac{\Omega'(p)}{\Omega^2}L^s\} \\ &= \frac{D_p(p+t,x^s) + D_p(p+t,x^u) - Y_p + Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial p} + Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial p}}{-Y_p + Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial p} + Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial p}} \\ &\quad\times\left\{-\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s\right\} \ .\end{aligned} \tag{23}$$

Finally, substituting the above equation for Δ, equation (21) becomes:

$$\begin{aligned}\left.\frac{\partial \mathcal{L}}{\partial \zeta}\right|_{\zeta=0} &= \frac{-Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial \zeta} - Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial \zeta}}{-Y_p + Y_{\bar{l}^s}\frac{\partial l_q^s}{\partial p} + Y_{\bar{l}^u}\frac{\partial l_q^u}{\partial p}} \\ &\quad\times\left\{-\lambda_1 V_2^{su}\Omega'(p)L^u + \lambda_2 V_2^{us}\frac{\Omega'(p)}{\Omega^2}L^s\right\} \ .\end{aligned} \tag{24}$$

Therefore, we proved lemma 1.

Acknowledgments

I thank Charles Brown, Yan Chen, James Hines, Miles Kimball, Joel Slemrod and Michelle White at the University of Michigan, Eiji Tajika at Hitotsubashi University, Nicholas Stern (editor) and anonymous referees of the journal for comments and suggestions. I very much appreciate careful reading of the draft and comments by my adviser, Roger Gordon. The responsibility for all remaining errors is mine.

References

Atkinson, A.B and Stiglitz, J.E., 1976, "The design of tax structure: direct versus indirect taxation", *Journal of Public Economics*, 6, pp55-75.

Atkinson, A.B. and Stiglitz, J.E., 1980, *Lecture on Public Economics* , McGraw-Hill.

Boadway, R. and Keen, M., 1993, "Public goods, self-selection and optimal income taxation," *International Economic Review*, 34, pp463-478.

Boadway, R. and Marchand, M., 1995, "The use of public expenditure for redistribution of Income," *Oxford Economic Papers*, 47, pp45-59.

Christiansen, V.A., 1984, "Which commodity taxes should supplement the income tax?," *Journal of Public Economics*,24, pp195-220.

Diamond, P.A. and Mirrlees, J.A., 1971, "Optimal taxation and public production 1: production efficiency", *American Economics Review*, 61, pp8-27.

Feldstein, M., 1973, "On the optimal progressivity of the income tax," *Journal of Public Economics*, pp357–366.

Guesnerie, R. and Seade, J., 1982, "Nonlinear pricing in a finite economy," *Journal of Public Economics*, pp157-179.

Gibbons, R., 1992, *Game Theory for Applied Economist*, Princeton University Press.

Gordon, R.H., 1986, "Taxation of investment and saving in a world economy, *American Economic Review*, 76, pp 1086-1102.

Harberger, A., 1962, "The incidence of the corporation income tax," *Journal of Political Economy*,70, pp215-240.

Konishi, H., 1995, "A Pareto-improving commodity tax reform under a smooth non-linear income tax," *Journal of Public Economics*, pp 413-445.

Laffont, J. J. and Tirole, J., 1993, *A Theory of Incentives in Procurement and Regulation*, The MIT press.

Mirrlees, J.A.,1971, "An exploration in the theory of optimum income taxation," *Review of Economic*

Studies, 38, pp175-208.

Myerson, R., "Baysian equilibrium and incentive compatibility: an introduction" in *Social Goals and Social Organization* Hurwitz.L ed.

Naito, H., 1996, "Tariffs as a device to relax the incentive problem of progressive income tax system," *Research Seminar in International Economics Working Paper No.391.* , The University of Michigan.

Naito, H., 1996, "Pattern of trade distortion and capital market distortion in a Small Open Economy ," Manuscript, The University of Michigan.

Razin A. and Sadka, E., 1991, International Tax Competition and Gains from Tax Harmonization, *Economics-Letters*, 37(1), pp 69-76.

Stern, N., 1982, "Optimum taxation with Errors in administration, *Journal of Public Economics*, 17, 181-211.

Stiglitz, J.E., 1982, "Self-selection and pareto efficient taxation," *Journal of Public Economics*, 17, pp213-240.

Stolper, W. and Samuelson, P., 1941, "Protection and real wages," *Review of Economic Studies*, pp58-73.