

Dynamic Tax Reform and Second-Best Pareto Optima

Alan Krause^{*}

This version: July 15, 2004

Abstract

The tax reform literature, pioneered by Guesnerie [1977], uses static models but views tax reform as a dynamic process, i.e., as a policy-maker implementing incremental reforms over time. This paper studies tax reform in a dynamic version of the Diamond-Mirrlees-Guesnerie model in which agents leave bequests and the government can run (temporary) budget deficits and surpluses. The main idea is that a tax reform in one period will affect endowments, equilibrium, and welfare in subsequent periods; thus, the process of tax reform cannot be analyzed as a sequence of static economies. The paper undertakes a tax reform analysis *a la* Guesnerie, but with an added focus on welfare improving reforms for each generation. Second-best Pareto optima are then characterized, and these conditions are compared to the static optimal tax formulae derived in the literature. In particular, the key Diamond-Mirrlees result that production efficiency is desirable at second-best optima no longer holds in the presence of (effective) restrictions on the taxation of private savings. Restrictions on government savings (including balanced budget restrictions), however, do not disturb the desirability of production efficiency. Finally, the effects of certain intertemporal political constraints on the tax reform process are also considered.

JEL classification codes: D5, D6, H2.

Keywords: Dynamic tax reform, second-best Pareto optima, capital taxation.

Acknowledgements: I thank Dustin Chambers, Jang-Ting Guo, Sushama Murty, Prasanta Pattanaik, and especially Charles Blackorby and R. Robert Russell for many helpful discussions. Participants at UCR's Economic Theory Colloquia, the 2004 Meeting of the Canadian Economic Association, and the 2004 Australasian Meeting of the Econometric Society also provided valuable comments. Any remaining errors or shortcomings are my responsibility alone.

* Department of Economics, University of California (Riverside), CA. 92521.
E-mail: alankrause@hotmail.com, alan.krause@email.ucr.edu.

1. Introduction

The relevance of optimal tax formulae, as developed by Diamond and Mirrlees [1971] among others, has been questioned on several grounds. In particular, the implementation of such formulae would likely require a major upheaval of the existing tax system, which policy-makers are unwilling or unable to do. In practice, actual changes are ‘slow and piecemeal’. This has shifted the emphasis from optimal tax or ‘tax design’ to ‘tax reform’, see especially Guesnerie [1977, 1995]. The tax reform literature takes the existing tax system and its imperfections as given, and examines the conditions under which there exist (differential) changes in taxes that are equilibrium preserving and Pareto improving, i.e., feasible reforms that increase the welfare of every agent in the economy.¹

Despite using a static model, it is clear that Guesnerie [1977, 1995] views tax reform as a dynamic process, i.e., as a policy-maker implementing incremental reforms over time. The dynamic aspect is analyzed explicitly in Fogelman, Quinzii, and Guesnerie [1978], who focus on the standard technical questions raised by dynamic systems: existence and stability. The present paper, however, studies a specific aspect of the dynamic process, namely, the implications for tax reform of agents leaving bequests and the government running (temporary) budget deficits and surpluses. The main idea is fairly straightforward: a tax reform in one period will affect bequests and therefore endowments, equilibrium, and welfare in subsequent periods. Thus, the process of tax

¹ The optimal tax literature has also been criticized for its reliance on the existence of a social welfare function. Tax reform analysis does not require a social welfare function to be specified, although it can be helpful in some aspects.

reform cannot be analyzed as a sequence of static economies; instead, the economies are dynamically linked. Accordingly, the taxation of capital plays an important role, and the government can transfer resources through time itself by issuing government bonds. These policy instruments are not available to the government in static models, in which there are no savings and the government's budget must always be balanced.

The paper employs a simple three-period model which is sufficient to capture the phenomena we are interested in, and the finite horizon prevents the government from running a Ponzi scheme. If the government undertakes a policy reform in period 1 bequests are likely to change, and therefore endowments, equilibrium, and welfare in subsequent periods will be disturbed. Thus, the government may be obligated to undertake reforms in periods 2 and 3, at least to restore equilibrium and balance its budget. Similarly, if the government does not reform in period 1 but does so in period 2, equilibrium and welfare in period 3 will be disturbed via the bequest and/or government bond mechanism. In sum, the government must choose a policy reform process rather than a sequence of independent reforms. The model is described in greater detail in Section 2.

Section 3 undertakes a tax reform analysis *a la* Guesnerie [1977, 1995]. That is, we address the central questions of tax reform analysis: (i) does there exist an equilibrium preserving reform that makes a particular generation better-off?, and the usual question addressed in the tax reform literature: (ii) does there exist an equilibrium preserving reform that makes every agent (generation) better-off? This provides the basis for a

characterization of second-best Pareto optima in Section 4, i.e., conditions under which there does not exist a feasible reform that makes every generation better-off are derived.

Section 5 examines the implications of certain constraints on the policy reform process which have received a great deal of attention both in academic circles and in public policy debates. The first type of constraint considered is restrictions on the taxation of capital and the government's ability to issue debt (say because the legislature will not approve such changes). The second type of constraint considered is a political constraint — that the welfare of generation i cannot be reduced in the process of implementing a social-welfare improving reform (say by the want of the government to get re-elected). The main finding is that the key Diamond-Mirrlees result that production efficiency is desirable at second-best optima no longer holds in the presence of (effective) restrictions on the taxation of private savings. Restrictions on government savings (including period-by-period balanced budget restrictions), however, do not disturb the desirability of production efficiency. Regarding the political constraint, it is shown to have a surprisingly straightforward effect on the conditions for second-best optimality. Section 6 contains some concluding remarks and summarizes the main results of the paper.

2. The Model

The model has three periods, with consumers living for one period. There is a single representative consumer in each period, so the terms 'consumer' and 'generation' are used interchangeably. The dynamic structure of the model is broadly similar to the finite horizon OLG model of Blackorby and Brett [1999, 2004], but the key feature of the

present model is that each generation cares about its own welfare as well as that of the following generation. This altruism materializes as a bequest. There are two types of bequest. The first is a storable capital good which forms the basis for the capital stock and can be purchased from the firm. The second is a bond which can be purchased from the government. From the point of view of each generation, the capital good and government bond are perfect substitutes, as both shift-out the budget constraint of the recipient generation. Competition between the government and firm for private savings ensures that the prices of the capital good and government bond are the same. The details of the model are now described.

2.1. Consumers

Generation 1 chooses its net (of endowment) consumption vector of non-storable goods $x^1 \in \mathbb{R}^n$, its leisure time $l^1 \in [0, 1]$, and its bequest $\kappa^1 := k^1 + b^1 \in \mathbb{R}_+$ to solve the following program:

$$V^1(\pi^1, \rho^1, \omega^1, R) = \max U^1(x^1, l^1, \kappa^1)$$

$$\text{subject to: } \pi^1 x^1 + \rho^1 \kappa^1 \leq \omega^1 (1 - l^1) + R \quad (2.1)$$

where superscripts indicate the time period and corresponding generation, $V(\cdot)$ is the indirect utility function, $\pi \in \mathbb{R}_+^n$ is the consumer price vector corresponding to the net consumption vector x , $\rho \in \mathbb{R}_+$ is the consumer price of the capital good $k \in \mathbb{R}_+$ and government bond $b \in \mathbb{R}$ (which is the mechanism for leaving bequests), $\omega \in \mathbb{R}_+$ is the wage (or consumer price of leisure), where the endowment of leisure is normalized to unity so that $(1 - l)$ is labor supplied, $R \in \mathbb{R}$ is a lump-sum tax, which is restricted to be

the same for each generation,² and $U(\cdot)$ is the direct utility function. The restriction that $\kappa^1 \in \mathbb{R}_+$ ensures that generation 1 cannot die in debt. However, the purchase of government bonds can be positive or negative, depending upon whether the government is running a budget deficit or surplus.

The inclusion of the bequest in generation 1's direct utility function is simpler (and arguably more realistic) than the usual assumption made regarding bequests in OLG models — that each generation knows the utility functions of the following generations and acts as if maximizing an infinite horizon utility function.³ Program (2.1) simply says that generation 1 divides its income between purchases of non-storable goods for itself and a bequest for generation 2 to maximize a utility function which represents generation 1's preferences over its own consumption and (in some sense) that of generation 2. In this framework, bequests provide a *positive externality* to the recipient generation, as they boost the recipient's income and indirect utility (see below).

Similarly, generation 2 chooses its net consumption vector $x^2 \in \mathbb{R}^n$, its leisure time $l^2 \in [0, 1]$, and its bequest $\kappa^2 \in \mathbb{R}_+$ to solve the following program:

$$V^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) = \max U^2(x^2, l^2, \kappa^2)$$

$$\text{subject to: } \pi^2 x^2 + \rho^2 \kappa^2 \leq \omega^2(1 - l^2) + R + \sigma^2 \kappa^1 \quad (2.2)$$

² R is sometimes called a poll tax or demogrant. The restriction that R be the same for each generation makes the model second-best, as the government does not have access to personalized lump-sum transfers.

³ See chapter 3 in Blanchard and Fischer [1989]. However, Hu [1979] for example also includes bequests in the direct utility function. This has been referred to as the 'joy of giving' motive for leaving bequests, e.g., like the joy of giving birthday presents.

where $\sigma^2 \in \mathbb{R}_+$ is the price generation 2 receives from selling the bequest it obtains from generation 1 to the firm (see below) and government at the beginning of the second period.

Generation 3 does not leave a bequest, as it is the last generation. Alternatively, generation 3 can be thought of as a purely non-altruistic generation. It chooses its net consumption vector $x^3 \in \mathbb{R}^n$ and leisure time $l^3 \in [0, 1]$ to solve the following program:

$$V^3(\pi^3, \omega^3, R, \sigma^3, \kappa^2) = \max U^3(x^3, l^3)$$

$$\text{subject to: } \pi^3 x^3 \leq \omega^3(1 - l^3) + R + \sigma^3 \kappa^2 \quad (2.3)$$

Note that generation 1 can be thought of as the ‘start-up’ generation, as it does not receive a bequest from the (non-existent) previous generation. Similarly, generation 3 is the ‘shut-down’ generation, as it does not leave a bequest. On the other hand generation 2, which both receives and leaves bequests, is of the usual type studied in infinite horizon models. Also note that it is implicit in the above formulation that the government taxes away all pure profits, as the consumers have no profit income. The assumption of a 100 percent tax on profits, or constant returns to scale in production which implies zero maximal profits, is standard in the optimal tax and tax reform literatures.

2.2. *Prices and Taxes*

Consumer prices are disconnected from producer prices by a complete system of specific taxes. That is:

$$\pi^1 = p^1 + t_x^1 \quad \pi^2 = p^2 + t_x^2 \quad \pi^3 = p^3 + t_x^3$$

$$\begin{aligned}
\rho^1 &= r^1 + t_k^1 & \rho^2 &= r^2 + t_k^2 & & (2.4) \\
\omega^1 &= w^1 + t_l^1 & \omega^2 &= w^2 + t_l^2 & \omega^3 &= w^3 + t_l^3 \\
& & \sigma^2 &= s^2 + t^2 & \sigma^3 &= s^3 + t^3
\end{aligned}$$

where $p \in \mathbb{R}_+^n$ is the producer price vector corresponding to x , $r \in \mathbb{R}_+$ is the price the firm receives from selling k , $w \in \mathbb{R}_+$ is the price the firm pays for hiring labor, and $s \in \mathbb{R}_+$ is the price the firm pays for purchases of k . Taxes (or subsidies) are simply the difference between consumer prices and producer prices.

2.3. Production

The supply side of the economy consists of a single, aggregate, profit maximizing firm whose technology changes over time, making this formulation consistent with any assumptions regarding capital depreciation or technical progress. The firm's profit functions are given by $\psi_1(p^1, r^1, w^1)$, $\psi_2(p^2, r^2, w^2, s^2)$, and $\psi_3(p^3, w^3, s^3)$. Application of Hotelling's Theorem to these profit functions yields the firm's supply and demand functions:

$$\begin{aligned}
x_1 &= x_1(p^1, r^1, w^1) & x_2 &= x_2(p^2, r^2, w^2, s^2) & x_3 &= x_3(p^3, w^3, s^3) \\
k_1 &= k_1(p^1, r^1, w^1) & k_2 &= k_2(p^2, r^2, w^2, s^2) & & (2.5) \\
L_1 &= L_1(p^1, r^1, w^1) & L_2 &= L_2(p^2, r^2, w^2, s^2) & L_3 &= L_3(p^3, w^3, s^3) \\
& & K_2 &= K_2(p^2, r^2, w^2, s^2) & K_3 &= K_3(p^3, w^3, s^3)
\end{aligned}$$

where the upper-case L and K are used to denote the firm's demands for labor and capital, i.e., it purchases capital from the consumers at the beginning of periods 2 and 3. The firm does not purchase any capital in period 1, as there is no supply by consumers, nor does it

produce any capital in period 3, as there is no demand by consumers. Note that subscript notation indicates supply or demand by the firm.

2.4. General Competitive Equilibrium

Equilibrium is obtained if and only if:

Period 1:

$$x^1(\pi^1, \rho^1, \omega^1, R) - x_1(p^1, r^1, w^1) \leq 0^{(n)} \quad (2.6)$$

$$k^1(\pi^1, \rho^1, \omega^1, R) - k_1(p^1, r^1, w^1) \leq 0 \quad (2.7)$$

$$b^1(\pi^1, \rho^1, \omega^1, R) - b_1 \leq 0 \quad (2.8)$$

$$L_1(p^1, r^1, w^1) + l^1(\pi^1, \rho^1, \omega^1, R) - 1 \leq 0 \quad (2.9)$$

Period 2:

$$x^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) - x_2(p^2, r^2, w^2, s^2) \leq 0^{(n)} \quad (2.10)$$

$$k^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) - k_2(p^2, r^2, w^2, s^2) \leq 0 \quad (2.11)$$

$$b^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) - b_2 \leq 0 \quad (2.12)$$

$$L_2(p^2, r^2, w^2, s^2) + l^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) - 1 \leq 0 \quad (2.13)$$

$$K_2(p^2, r^2, w^2, s^2) - k^1(\pi^1, \rho^1, \omega^1, R) \leq 0 \quad (2.14)$$

$$B_2 - b^1(\pi^1, \rho^1, \omega^1, R) \leq 0 \quad (2.15)$$

Period 3:

$$x^3(\pi^3, \omega^3, R, \sigma^3, \kappa^2) - x_3(p^3, w^3, s^3) \leq 0^{(n)} \quad (2.16)$$

$$L_3(p^3, w^3, s^3) + l^3(\pi^3, \omega^3, R, \sigma^3, \kappa^2) - 1 \leq 0 \quad (2.17)$$

$$K_3(p^3, w^3, s^3) - k^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) \leq 0 \quad (2.18)$$

$$B_3 - b^2(\pi^2, \rho^2, \omega^2, R, \sigma^2, \kappa^1) \leq 0 \quad (2.19)$$

where b_1 and b_2 denote the supply of bonds by the government in periods 1 and 2, and B_2 and B_3 denote the government's demand (repurchases) of bonds in periods 2 and 3. The model's finite horizon prevents the government from engaging in some sort of Ponzi game, with Walras' Law ensuring that the government's budget is balanced over the three periods. Note that the periods are linked by the demand for and supply of capital and government bonds. Thus, reforms in periods 1 and 2 will disturb equilibrium in periods 2 and 3, i.e., we do not have just a sequence of static economies. For analytical purposes, it is assumed that the status quo is a 'tight' equilibrium, i.e., equations (2.6) – (2.19) all hold with equality.

3. Characterizing Reform Possibilities

Apart from the fixed 100 percent tax on pure profits, the government's policy instruments are the taxes on commodities, capital, and labor, as well as the bonds and demogrant. As the government does not supply nor demand commodities, it uses its instruments only to redistribute income (or to pick equilibria in the sense of the Second Theorem of Welfare Economics), not to finance public goods and services.

A *policy reform process* can be defined as a vector $dP := \langle dP^1, dP^2, dP^3, dR \rangle$

where:

$$dP^1 := \langle dp^1, dt_x^1, dr^1, dt_k^1, dw^1, dt_l^1, db_1 \rangle$$

corresponds to changes in the government's first-period instruments, and similarly:

$$dP^2 := \langle dp^2, dt_x^2, dr^2, dt_k^2, dw^2, dt_l^2, ds^2, d\tau^2, db_2, dB_2 \rangle$$

$$dP^3 := \langle dp^3, dt_x^3, dw^3, dt_l^3, ds^3, d\tau^3, dB_3 \rangle$$

correspond to changes in the government's second- and third-period instruments.

A reform dP satisfies the equilibrium conditions (2.6) – (2.19) if and only if $\nabla Z dP \leq 0^{(3n+11)}$, where ∇Z is a Jacobian matrix of excess-demand derivatives and is defined explicitly in the Appendix. The welfare of generation 1 increases if and only if $dV^1(\cdot) = \nabla V^1 dP > 0$, where ∇V^1 is the gradient of generation 1's indirect utility function (see the Appendix for details). Analogously, the welfares of generations 2 and 3 increase if and only if $dV^2(\cdot) = \nabla V^2 dP > 0$ and $dV^3(\cdot) = \nabla V^3 dP > 0$.

The above formulation gives the government $6n + 19$ instruments, but it must satisfy the $3n + 11$ equilibrium conditions in (2.6) – (2.19) which, by the Implicit Function Theorem, suggests $3n + 8$ degrees of freedom in picking equilibria. However, six of these degrees of freedom are only apparent since a consumer price and a producer price can be normalized to one in each period.⁴ This leaves the government with $3n + 2$ degrees of freedom in picking equilibria. We are interested in how these equilibria compare to the status quo on the basis of certain normative criteria, which is the issue now addressed.

3.1. *Equilibrium Preserving and Welfare Improving Reforms*

We now address the central questions of tax reform analysis by characterizing the possibility (or otherwise) of equilibrium preserving and welfare improving reforms. Specifically, we want to know if there exists a feasible reform that makes a particular

generation better-off, and if there exists a feasible reform that makes every generation better-off. The answer is given by Proposition 1 and its corollaries.

Proposition 1:

Let $\Gamma \subseteq \mathbb{R}^{6n+19}$ be the cone generated by taking non-negative linear combinations of the rows of ∇Z . If $\nabla V^i \in \Gamma$, then there does not exist an equilibrium preserving policy reform process that makes generation i better-off. If $\nabla V^i \notin \Gamma$, then there exists an equilibrium preserving policy reform process that makes generation i better-off.

Proof:

By Farkas' Theorem there exists a $\mu \in \mathbb{R}_+^{3n+11}$ such that $\mu \nabla Z = \nabla V^i$, or there exists a dP such that $\nabla Z dP \leq 0^{(3n+11)}$ and $\nabla V^i dP > 0$, but never both. Next, note that $\nabla V^i \in \Gamma$ implies that there must exist a $\mu \in \mathbb{R}_+^{3n+11}$ such that $\mu \nabla Z = \nabla V^i$, which by Farkas' Theorem implies no solution dP to $\nabla Z dP \leq 0^{(3n+11)}$ and $\nabla V^i dP > 0$. That is, there does not exist an equilibrium preserving policy reform process that makes generation i better-off. Conversely, $\nabla V^i \notin \Gamma$ implies that there cannot exist a $\mu \in \mathbb{R}_+^{3n+11}$ such that $\mu \nabla Z = \nabla V^i$, which means there exists a reform dP such that $\nabla Z dP \leq 0^{(3n+11)}$ and $\nabla V^i dP > 0$. ■

A discussion of Farkas' Theorem can be found in Mangasarian [1969]. Following the existing tax reform literature, we are particularly interested in the possibility of

⁴ This follows from the consumers' demand functions being homogenous of degree zero in consumer prices and the demogrant, and the producer's supply and demand functions being homogenous of degree zero in producer prices.

equilibrium preserving and Pareto improving reforms.⁵ In the present context, these are those (ideal) reforms which are feasible and improve the welfare of every generation.

Corollary 1.1:

If $\nabla V^1 \in \Gamma$ or $\nabla V^2 \in \Gamma$ or $\nabla V^3 \in \Gamma$, then there does not exist an equilibrium preserving and Pareto improving policy reform process.

Corollary 1.2:

Suppose $\Gamma \cap -\Gamma = \{0^{(6n+19)}\}$, i.e., Γ is a pointed cone. Let $-P(\Gamma)$ be the negative polar cone of Γ . If $-P(\Gamma) \subseteq -\Gamma$ and $\nabla V^1 \in -\Gamma$ and $\nabla V^2 \in -\Gamma$ and $\nabla V^3 \in -\Gamma$, then there exists an equilibrium preserving and Pareto improving policy reform process.

Corollary 1.3:

Suppose $\Gamma \cap -\Gamma = \{0^{(6n+19)}\}$, i.e., Γ is a pointed cone. Let $-P(\Gamma)$ be the negative polar cone of Γ . If $-\Gamma \subseteq -P(\Gamma)$ and $\nabla V^1 \in -P(\Gamma)$ and $\nabla V^2 \in -P(\Gamma)$ and $\nabla V^3 \in -P(\Gamma)$, then there exists an equilibrium preserving and Pareto improving policy reform process.

Corollary 1.4:

If $\nabla V^1 \notin \Gamma$ and $\nabla V^2 \notin \Gamma$ and $\nabla V^3 \notin \Gamma$, and at least one of these vectors is an element of the complement of $\Gamma \cup -P(\Gamma) \cup -\Gamma$ (relative to \mathbb{R}^{6n+19}), then there may or may not exist an equilibrium preserving and Pareto improving policy reform process.

Some intuition for Proposition 1 and its corollaries can be obtained by reference to Figures 1 and 2. Figure 1 illustrates an example (in a stylized two-dimensional manner) in which there exists an equilibrium preserving reform that makes generation 1

⁵ See Guesnerie [1977: Proposition 4], Guesnerie [1995: Theorem 1, p.145], Brett [1998: Proposition 2], and more recently Murty and Russell [2003: Theorem 1].

better-off, but generations 2 and 3 are made worse-off. A reform dP is equilibrium preserving if it forms an obtuse angle with all the rows of ∇Z . Letting Γ denote the cone generated by taking non-negative linear combinations of the rows of ∇Z , the set of equilibrium preserving reforms is therefore $-P(\Gamma)$, i.e., the negative polar cone of Γ . A reform dP makes generation i better-off if dP forms a strictly acute angle with the gradient of generation i 's indirect utility function ∇V^i . In Figure 1, the shaded area is the set of equilibrium preserving reforms which make generation 1 better-off. However, all equilibrium preserving reforms make generations 2 and 3 worse-off, as they all form obtuse angles with ∇V^2 and ∇V^3 .

Further insight into Proposition 1 can be obtained by considering the hypothetical problem of choosing the policy instruments P to maximize the welfare of generation i subject to the equilibrium preserving constraints (2.6) – (2.19). Suppose \tilde{P} solves this problem, then by the Kuhn-Tucker Theorem there exists a $\tilde{\mu} \in \mathbb{R}_+^{3n+1}$ such that:

$$\tilde{\mu} \nabla Z = \nabla V^i \tag{KT}$$

where $\tilde{\mu}$ is a vector of Kuhn-Tucker multipliers. If $\nabla V^i \notin \Gamma$, then there cannot exist a vector $\tilde{\mu} \in \mathbb{R}_+^{3n+1}$ such that (KT) is satisfied. That is, the status quo equilibrium does not ‘look like’ the policy instruments have already been chosen to maximize the welfare of generation i . In this case, there must exist a feasible reform which makes generation i better-off. On the other hand, if there does not exist a feasible reform which makes generation i better-off, then (KT) is satisfied and the status quo equilibrium ‘looks like’ the policy instruments have already been chosen to maximize the welfare of generation i .

In order for there to exist an equilibrium preserving and Pareto improving reform, the vectors ∇V^1 , ∇V^2 , and ∇V^3 must be ‘close enough’ in the sense made precise by Corollaries 1.1 – 1.4. Figure 2 illustrates an example (again in a stylized two-dimensional manner) in which there exists an equilibrium preserving and Pareto improving reform. In this example ∇V^1 , ∇V^2 , and ∇V^3 are all elements of $-\Gamma$, which is a sufficient condition for there to exist an equilibrium preserving and Pareto improving reform (refer Corollary 1.2).

While Proposition 1 might be criticized for being too abstract, this is unavoidable. The question of tax reform is a difficult one, and it does not lend itself to a simple answer with a straightforward economic interpretation.⁶ However, the usefulness of Proposition 1 is that it forms the basis for the characterization of second-best Pareto optima (see below), and it provides an empirically implementable methodology for identifying the possibility of welfare improving reforms. That is, the criteria of Proposition 1 can be checked empirically by constructing the cone Γ and with knowledge of the gradients of the consumers’ indirect utility functions. The construction of Γ requires estimates of various aggregate demand and supply derivatives. The gradients of the consumers’ indirect utility functions are not directly observable. However, by Roy’s Theorem, ∇V^i is proportional to consumer i ’s consumption vector, and survey data are available which provide details on household consumption (e.g., the household survey used to construct the CPI). In sum, the information requirements of Proposition 1 (and its corollaries) can

⁶ Unlike optimal tax, tax reform analysis is not fundamentally an optimization exercise and therefore cannot be studied using well known and well established optimization techniques.

be met. Examples of empirical applications of tax reform theory, following the work of Guesnerie [1977], are King [1983] and Ahmad and Stern [1984].

4. Second-Best Pareto Optima

As noted by Guesnerie [1977] and Weymark [1979], necessary conditions for the non-existence of an equilibrium preserving and Pareto improving reform are, in fact, necessary conditions for the initial position to be Pareto optimal. In other words, these conditions characterize second-best optima as in the optimal tax literature, and are derived in the following proposition.

Proposition 2:

Let ∇V be the $3 \times (6n + 19)$ matrix formed by the vectors ∇V^1 , ∇V^2 , and ∇V^3 . If there does not exist a policy reform process dP such that:

$$\nabla Z dP \leq 0^{(3n+11)}$$

$$\nabla V dP \gg 0^{(3)}$$

That is, if there does not exist an equilibrium preserving and Pareto improving policy reform process, then the economy is at a local second-best optimum and there exist two vectors of multipliers $\lambda > 0^{(3)}$ and $\mu \geq 0^{(3n+11)}$ such that:

$$\lambda \nabla V = \mu \nabla Z \tag{4.1}$$

Proof:

Follows from an application of Motzkin's Theorem of the Alternative. ■

A discussion of Motzkin's Theorem can be found in Mangasarian [1969]. The existence of the multipliers guaranteed by Proposition 2 is not very informative in itself.

However, each multiplier λ_i can be interpreted as the *implied welfare weight* of generation i , and the multipliers μ are the *social shadow prices* of the $3n + 11$ commodities. That is, if the government could choose its policy instruments to maximize a social welfare function $W(V^1, V^2, V^3)$ subject to (2.6) – (2.19), then $\lambda_i = \partial W(\cdot) / \partial V^i$ and μ is the vector of multipliers attached to the $3n + 11$ constraints in (2.6) – (2.19). In other words, if there does not exist an equilibrium preserving and Pareto improving policy reform process, the economy ‘looks like’ the government has already chosen its policy instruments to maximize a social welfare function. For future reference, we define:

$$\beta_i = \frac{\partial W(\cdot)}{\partial V^i(\cdot)} \frac{\partial V^i(\cdot)}{\partial I^i} = \lambda_i \frac{\partial V^i(\cdot)}{\partial I^i}$$

which is known in the optimal tax literature as the *social marginal utility of income* of generation i ($\partial V^i(\cdot) / \partial I^i$ is generation i ’s private marginal utility of income).

Expanding (4.1) yields the first-order (necessary) conditions for second-best optimality. Each equation in (4.1) has a standard interpretation in the optimal tax literature: at an optimum, the *marginal social benefit* of a change in any instrument is equal to the *marginal social cost* of that change.⁷ For example, a tax decrease which reduces the consumer price of some good will boost consumption of that good and welfare. This is the benefit. The cost is that the increase in demand must be met by an increase in supply by transferring resources from other sectors of the economy. Some characteristics of the second-best Pareto optima — in particular, how they compare to the

⁷ See, for example, Diamond and Mirrlees [1971: Equations 19 and 66], Dixit [1979: Equations 11 and 12], Guesnerie [1995: Theorem 8, p.178], and Blackorby and Brett [2000: p.305].

optima of the static Diamond-Mirrlees-Guesnerie model — are discussed in the following corollaries. Proofs, where necessary, are relegated to the Appendix.

Corollary 2.1:

At all second-best Pareto optima production efficiency holds. That is, social shadow prices are proportional to producer prices in each period.

The term ‘production efficiency’ has been used in various guises in the optimal tax and tax reform literatures. Diamond and Mirrlees [1971] say production is efficient if technical rates of substitution in private and public sector production are equal. This is the usual ‘textbook’ definition of production efficiency. In Guesnerie [1977, 1995: chapter 3] an equilibrium is temporarily (production) inefficient if it involves excess supply of some goods, i.e., ‘non-tight’ equilibria. Blackorby and Brett [2000], see also Guesnerie [1995: chapter 4], define an equilibrium to be production efficient if producer prices are proportional to social shadow prices. This is equivalent to the Diamond-Mirrlees definition if there is a public producer, because social shadow prices are the relevant prices for public production decisions. It is in this latter sense that this paper refers to production efficiency.⁸

As noted by Guesnerie [1995: p.192], it is generally difficult to assess the relationship between social values and market prices (or more generally with technological conditions prevailing in the economy) in a second-best model with consumer prices disconnected from producer prices. However, Corollary 2.1 shows that social shadow prices are proportional to producer prices at all second-best optima. This

⁸ In order to minimize notation, the government’s provision of public goods is not explicitly modeled.

reinforces the key Diamond-Mirrlees result (but in a dynamic model with bequests which generate a positive externality) that production efficiency is desirable at second-best Pareto optima, despite the second-best nature of the model.⁹

Corollary 2.2:

At all second-best Pareto optima, the multipliers attached to the bond market equations (2.8), (2.12), (2.15), and (2.19) are all equal to zero. That is, differential changes in the government's debt instruments b_1 , b_2 , B_2 , and B_3 have no social value.

Corollary 2.2 simply follows from the fact that the government has direct control over its bond issues. If changes in its debt instruments can increase social welfare, then the economy cannot be at an optimum.

Corollary 2.3:

At all second-best Pareto optima, the marginal social benefits of increasing the taxes t_x^1 , t_x^2 , and t_x^3 on the first-, second-, and third-period consumption goods are given by, respectively:

$$\begin{aligned}
 & -\beta_1 x^1 + \beta_2 \sigma^2 \nabla_{\pi^1} \kappa^1 + \beta_3 \sigma^3 \nabla_{\pi^1} \kappa^2 \\
 & -\beta_2 x^2 + \beta_3 \sigma^3 \nabla_{\pi^2} \kappa^2 \\
 & -\beta_3 x^3
 \end{aligned}$$

⁹ Recall the Lipsey and Lancaster [1956] general theorem of the second best, which asserts that if a distortion prevents the attainment of one of the conditions for first-best Pareto optimality, then the other conditions, although attainable, are generally no longer desirable. The distortion in the present context is the commodity taxes, which prevents marginal rates of substitution in consumption from being equated to marginal rates of transformation in production.

Therefore, given sufficient substitutability between bequests and the consumption goods, the marginal social benefits of increasing t_x^1 and t_x^2 may be positive.

In static models, the marginal social benefit of increasing consumption taxes is necessarily negative. In the present dynamic model, however, generations 1 and 2 may substitute bequests for their own consumption, boosting the incomes and welfare of the recipient generations. For example, consider the first equation in Corollary 2.3. The first term shows the negative welfare effect of increasing t_x^1 on generation 1, which operates through higher first-period consumer prices. The second and third terms show the welfare effects on generations 2 and 3, which are positive if $\nabla_{\pi^i} \kappa^1$ and $\nabla_{\pi^i} \kappa^2$ are positive. Thus, it is possible that the marginal social benefit of increasing consumption taxes is positive, especially in regions of the Pareto frontier in which the implied welfare weight of the saving (i.e., bequesting) generation is low relative to the recipient's implied welfare weight (as governed by the magnitude of β_i).

The marginal social benefits of decreasing the taxes on labor income, however, are always positive as in static models, provided the bequests are normal goods (which is probably a safe assumption). That is:

Corollary 2.4:

At all second-best Pareto optima, the marginal social benefits of decreasing the taxes t_l^1 , t_l^2 , and t_l^3 on first-, second-, and third-period labor income are given by, respectively:

$$\beta_1(1-l^1) + \beta_2\sigma^2\nabla_{\omega^1}\kappa^1 + \beta_3\sigma^3\nabla_{\omega^1}\kappa^2$$

$$\beta_2(1-l^2) + \beta_3\sigma^3\nabla_{\omega^2}\kappa^2$$

$$\beta_3(1-l^3)$$

Therefore, if the bequests κ^1 and κ^2 are both normal goods, then the marginal social benefits of decreasing the taxes on labor income are always positive.

Consider the first equation in Corollary 2.4. The first term shows the direct benefit to generation 1 of a decrease in its labor income tax t_l^1 . The second term shows the benefit to generation 2 if generation 1 spends part of its increase in post-tax labor income on the bequest κ^1 . This increases generation 2's income, so the third term shows the benefit to generation 3 if part of the increase in generation 2's income is spent on the bequest κ^2 .

5. Constraints on the Policy Reform Process

It has thus far been assumed that the government can change all taxes and issue bonds at will. In reality, however, it is well known that no central authority controls all taxes, and typically the government must seek legislative approval for changes. It is also often the case that the government may be unable or unwilling to implement reforms which reduce the welfare of the current generation (say by the want to get re-elected), even if such reforms increase the welfare of future generations, i.e., even if the short-run costs are outweighed by the long-run benefits. Given that a key distinction between the optimal tax and tax reform approaches is that tax reform explicitly acknowledges the existing tax system and its imperfections, this section examines the implications of certain constraints on the policy reform process.

5.1. Restrictions on the Government's Intertemporal Policy Instruments

The taxation of capital has long been a controversial issue, and proposed increases in capital taxation may therefore be blocked by the legislature.¹⁰ Restrictions on the government's ability to issue debt, such as period-by-period balanced budget restrictions, have also been proposed. We explore the consequences of such restrictions on the policy reform process.

First, suppose the government is not permitted to increase the taxes on savings (i.e., purchases of the capital good) by generations 1 and 2. That is, we impose the restrictions $dt_k^1 \leq 0$ and $dt_k^2 \leq 0$. By Motzkin's Theorem, if there is no solution dP to:

$$\nabla Z dP \leq 0^{(3n+11)} \quad \nabla V dP \gg 0^{(3)} \quad \begin{pmatrix} 0^{(2n+1)} & -1 & 0^{(4n+17)} \\ 0^{(4n+6)} & -1 & 0^{(2n+12)} \end{pmatrix} dP \geq 0^{(2)}$$

That is, if there does not exist an equilibrium preserving and Pareto improving policy reform process (with the restrictions on capital taxation), then there exist vectors $\lambda > 0^{(3)}$, $\mu \geq 0^{(3n+11)}$, and $\theta \geq 0^{(2)}$ such that:

$$\lambda \nabla V + \theta \begin{pmatrix} 0^{(2n+1)} & -1 & 0^{(4n+17)} \\ 0^{(4n+6)} & -1 & 0^{(2n+12)} \end{pmatrix} = \mu \nabla Z \quad (5.1)$$

The system (5.1) is the same as the system (4.1) which characterizes unrestricted second-best optima, except the 4th and 11th equations in (5.1) include the multipliers $\langle \theta_1, \theta_2 \rangle$ which correspond to the restrictions on capital taxation. Note that if either of the restrictions are binding, i.e., if there exists an equilibrium preserving and Pareto

¹⁰ Blackorby and Brett [2004] address the question as to whether it is optimal to tax capital and summarize the existing literature.

improving reform in the absence of the restrictions, then $\theta \neq 0^{(2)}$. This leads to the following proposition, which is proved in the Appendix.

Proposition 3:

Suppose the government is not permitted to increase the taxes on savings by generations 1 and 2. Then production efficiency holds in period 1 (period 2) at all second-best Pareto optima if and only if the constraint on the taxation of savings in period 1 (period 2) is not binding. (Note: production efficiency always remains desirable in period 3.)

In other words, production efficiency is no longer a necessary condition for second-best optimality if the government cannot increase the taxes on capital. In regions of the Pareto frontier in which the constraints are not binding, production efficiency remains desirable. However, when one of the constraints is binding, the marginal social benefit of increasing the corresponding tax is greater than the marginal social cost. In this case, the government would like to increase the tax to directly determine bequests and the supply of capital available to the firm. Instead, it must distort producer prices relative to social shadow prices as an indirect (and imperfect) means of control. For example, the government might instruct a public producer to use more than the efficient quantity of capital in order to (indirectly) increase the price of capital. Roughly speaking, this is the classic ‘two wrongs make a right’ result of second-best theory.

The second way that the government can influence intertemporal income transfers is by issuing bonds itself. However, history suggests that government deficits and debt levels can become excessive, imposing a burden on future generations. In response, some

economists have argued in favor of various forms of fiscal constraint, the most extreme of which are balanced budget restrictions. The following result, which is proved in the Appendix, shows that such restrictions do not disturb production efficiency.

Proposition 4:

Production efficiency remains desirable in each period at all second-best Pareto optima in the presence of constraints on the government's debt instruments, including period-by-period balanced budget restrictions, whether or not the constraints are binding.

From the point of view of production efficiency, then, restrictions on the government's ability to issue debt are preferable to restrictions on the government's ability to tax private capital. Even though the capital good and government bonds are perfect substitutes (from the perspective of each generation), purchases and sales of bonds are transactions between consumers and the government only; they do not involve the production sector. The producer is affected only indirectly by the prices it receives and pays for capital.

On the issue of production efficiency, this paper contributes to a continuing literature which concludes that the Diamond-Mirrlees result is the exception rather than the rule. In the context of the issue of fiscal federalism, Blackorby and Brett [2000] show that production inefficiency is a characteristic of almost all (second-best) Pareto optima when the states control some taxes and the federal government controls others. Similarly, in an international setting, Keen and Wildasin [2004] show that when each country faces its own government's budget constraint the Diamond-Mirrlees result no longer applies. In

the dynamic model of the present paper, where the government can effect intertemporal transfers by influencing the savings behavior of the private sector and by issuing bonds itself, we show that (effective) restrictions on the taxation of private savings entail production inefficiency at second-best optima. However, restrictions on government savings do not disturb the desirability of production efficiency.

5.2. *Social Welfare and Political Constraints*

Following the existing tax reform literature, attention has focused thus far on the possibility of Pareto improving reforms, i.e., reforms which increase the welfare of every generation. However, the Pareto criterion can be criticized as being much too demanding, especially if the status quo equilibrium involves a highly unequal distribution of income. Moreover, in general, the set of feasible reforms which increase the welfare of *every* agent in the economy is likely to be small, if not empty. For these reasons, the government may instead refer to a social welfare function $W(V^1, V^2, V^3)$. However, the government may be unable or unwilling, for political reasons, to implement social-welfare improving reforms which decrease the welfare of, say, the current generation. This places a political constraint on the set of social-welfare improving reforms.

First, we show in Proposition 5 that a characterization analogous to that in Proposition 1 goes through using the social-welfare criterion (excluding for the moment political constraints). Note that a reform dP increases social welfare if and only if $dW(\cdot) = \nabla W dP > 0$, where ∇W is the gradient of the social welfare function and is defined explicitly in the Appendix.

Proposition 5:

If $\nabla W \in \Gamma$, then there does not exist an equilibrium preserving policy reform process that increases social welfare. If $\nabla W \notin \Gamma$, then there exists an equilibrium preserving policy reform process that increases social welfare.

Proof:

By Farkas' Theorem, there exists a $\mu \in \mathbb{R}_+^{3n+1}$ such that $\mu \nabla Z = \nabla W$, or there exists a dP such that $\nabla Z dP \leq 0^{(3n+1)}$ and $\nabla W dP > 0$, but never both. Next, note that $\nabla W \in \Gamma$ implies that there must exist a $\mu \in \mathbb{R}_+^{3n+1}$ such that $\mu \nabla Z = \nabla W$. Conversely, $\nabla W \notin \Gamma$ implies that there cannot exist a $\mu \in \mathbb{R}_+^{3n+1}$ such that $\mu \nabla Z = \nabla W$. ■

Corollary 5.1:

If $\nabla W \in \Gamma$ the economy is at a local second-best optimum and there exists a vector of multipliers $\mu \in \mathbb{R}_+^{3n+1}$ such that:

$$\nabla W = \mu \nabla Z \tag{5.2}$$

The system of equations (5.2) is *identical* to the system (4.1). Thus, it makes no difference whether second-best optima are characterized using the Pareto criterion or the social-welfare criterion. Intuition for this result is provided in the discussion of Proposition 6 below.

We can now examine the implications of political constraints on the policy reform process. Specifically, suppose the welfare of generation i cannot be reduced in the process of implementing a social-welfare improving reform. By Motzkin's Theorem, if there does not exist a policy reform process dP such that:

$$\nabla Z dP \leq 0^{(3n+1)} \quad \nabla W dP > 0 \quad \nabla V^i dP \geq 0$$

then there exists $\gamma > 0$, $\bar{\mu} \geq 0^{(3n+1)}$, and $\bar{\theta} \geq 0$ such that:

$$\nabla W + \theta \nabla V^i = \mu \nabla Z$$

where $\theta = \bar{\theta}/\gamma$ and $\mu = \bar{\mu}/\gamma$. If the political constraint is binding ($\theta > 0$) there exists a social-welfare improving reform, but all such reforms make generation i worse-off. The following result can now be stated.

Proposition 6:

Suppose the government is prevented from implementing social-welfare improving reforms which make generation i worse-off. Then second-best Pareto optima are still characterized by (4.1), except λ_i is replaced by $\lambda_i + \theta$.

That is, the political constraint simply acts to increase the implied welfare weight of generation i from λ_i to $\lambda_i + \theta$. The intuition can be explained by reference to Figure 3. Conditions for second-best optimality hold at *all* points on the Pareto frontier. The political constraint, requiring that the welfare of generation 1 (for example) not be reduced below its initial level (say \bar{U}^1), means that the economy is initially at a point on the second-best Pareto frontier (say \bar{E}) which is relatively favorable to generation 1. In the absence of the political constraint, the government could implement a social-welfare improving reform and move the economy towards E^* .

The bottom line is that the social-welfare approach and political constraints do not fundamentally alter the necessary conditions for second-best Pareto optimality obtained in Section 4. This implies that the results obtained earlier in the paper are robust to this

alternative approach, with the simple exception that the implied welfare weight of generation i is increased if the political constraint favoring generation i is binding.

6. Concluding Comments

The existing tax reform literature views tax reform as a dynamic process, but it uses static models. Examples of this literature include: Guesnerie [1977, 1995], Diewert [1978], Dixit [1979], Weymark [1979], Brett [1998], and Murty and Russell [2003]. This paper has examined tax reform in a dynamic version of the Diamond-Mirrlees-Guesnerie model (which is the classic model used in optimal tax and tax reform), whereby the periods are dynamically linked by bequests and the government can run (temporary) budget deficits and surpluses. In this framework, the process of tax reform cannot be analyzed as a sequence of static economies. The main contributions of the paper are:

1. the development of a simple, tractable, dynamic model which makes it possible to analyze tax reform as a dynamic process;
2. the development of a tax reform methodology which makes it possible to derive the conditions under which each generation (or agent) can and cannot be made better-off (in addition to the usual Pareto improving conditions);
3. second-best Pareto optima have been characterized as in the optimal tax literature, with special attention paid to intertemporal conditions and how these compare to static optimal tax formulae;
4. production efficiency is shown to be desirable at all second-best Pareto optima, reinforcing the key Diamond-Mirrlees result (but in a dynamic model), but not if

- there are restrictions on the taxation of private savings. Restrictions on government savings, however, do not disturb the desirability of production efficiency; and
5. it makes no difference whether second-best optima are characterized using the Pareto criterion or the social-welfare criterion, and political constraints which favor some generation simply increase the implied welfare weight of that generation (without disturbing fundamentally the conditions for second-best optimality).

Finally, it has been suggested that dynamic tax reform could be studied in a model in which agents live for more than one period and save for life cycle reasons. While this is certainly feasible, it would slightly complicate the model and simply change the motive for saving (from an altruistic bequest motive to saving for their own future consumption). Moreover, the fundamental issues addressed in this paper would remain the same, i.e., the process of tax reform cannot be analyzed as a sequence of static economies as a reform in one period will affect savings and therefore endowments, equilibrium, and welfare in subsequent periods.

FIGURE 1

The Geometry of Proposition 1 and its Corollaries

There exists an equilibrium preserving reform that makes generation 1 better-off (shaded area), but generations 2 and 3 are made worse-off.

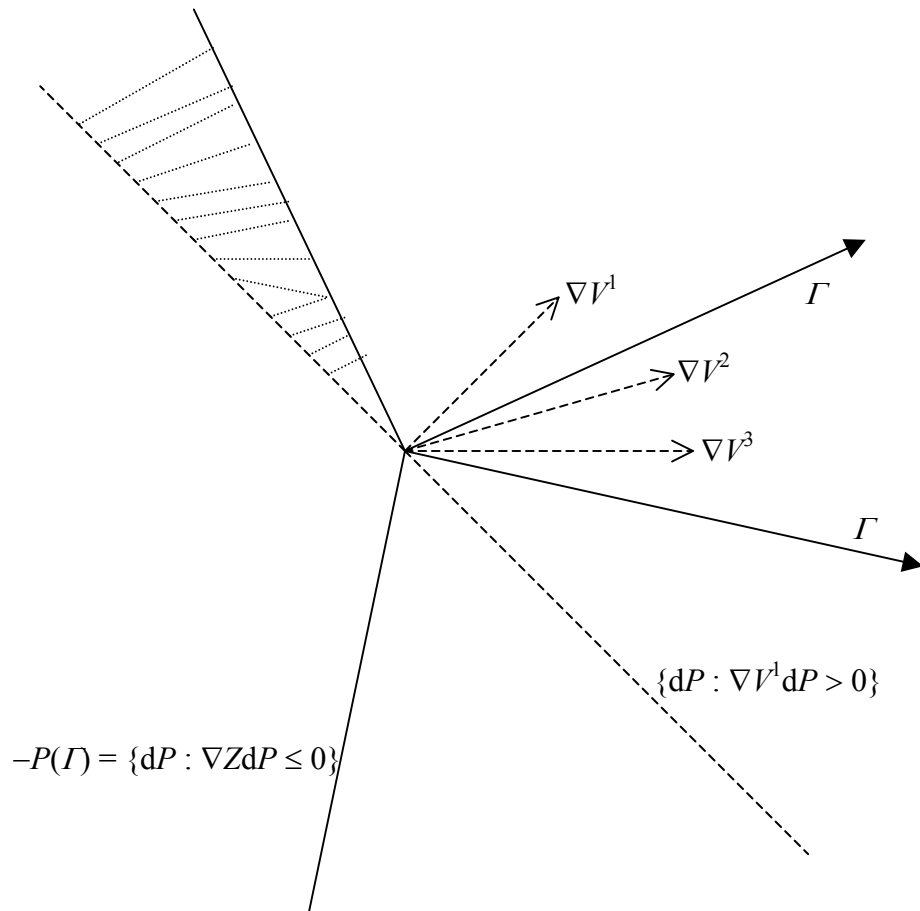


FIGURE 2

The Geometry of Proposition 1 and its Corollaries

There exists an equilibrium preserving and Pareto improving reform.

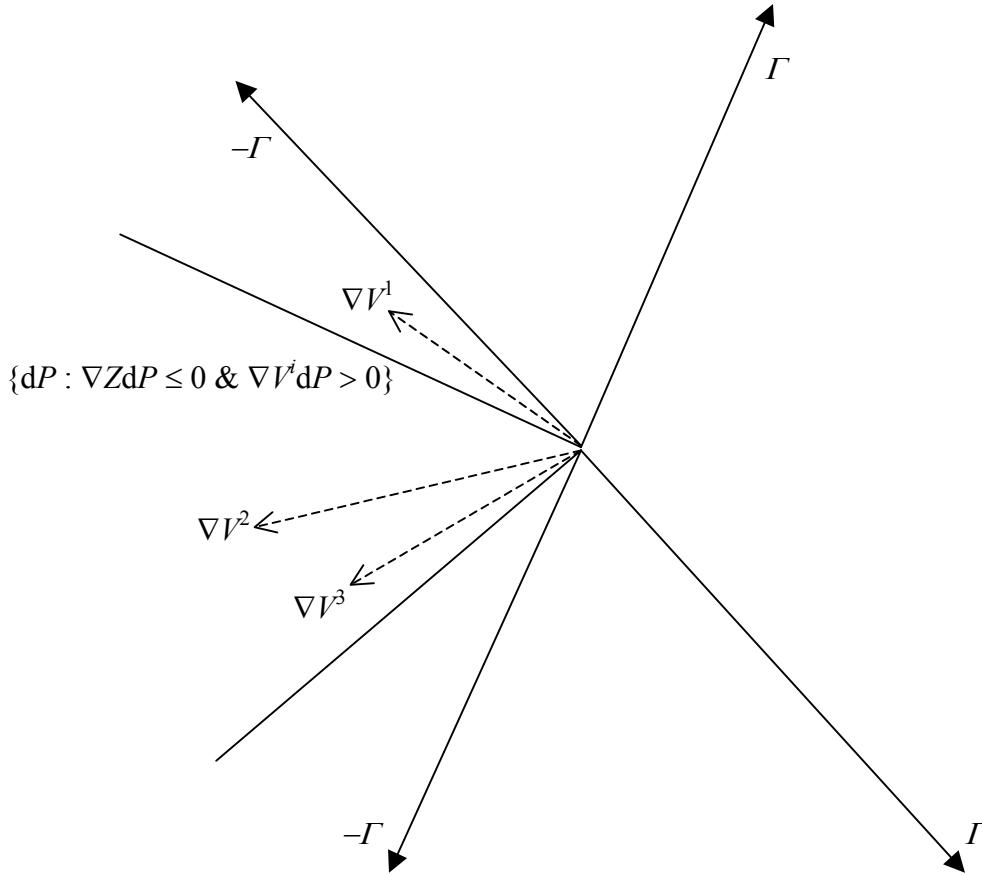
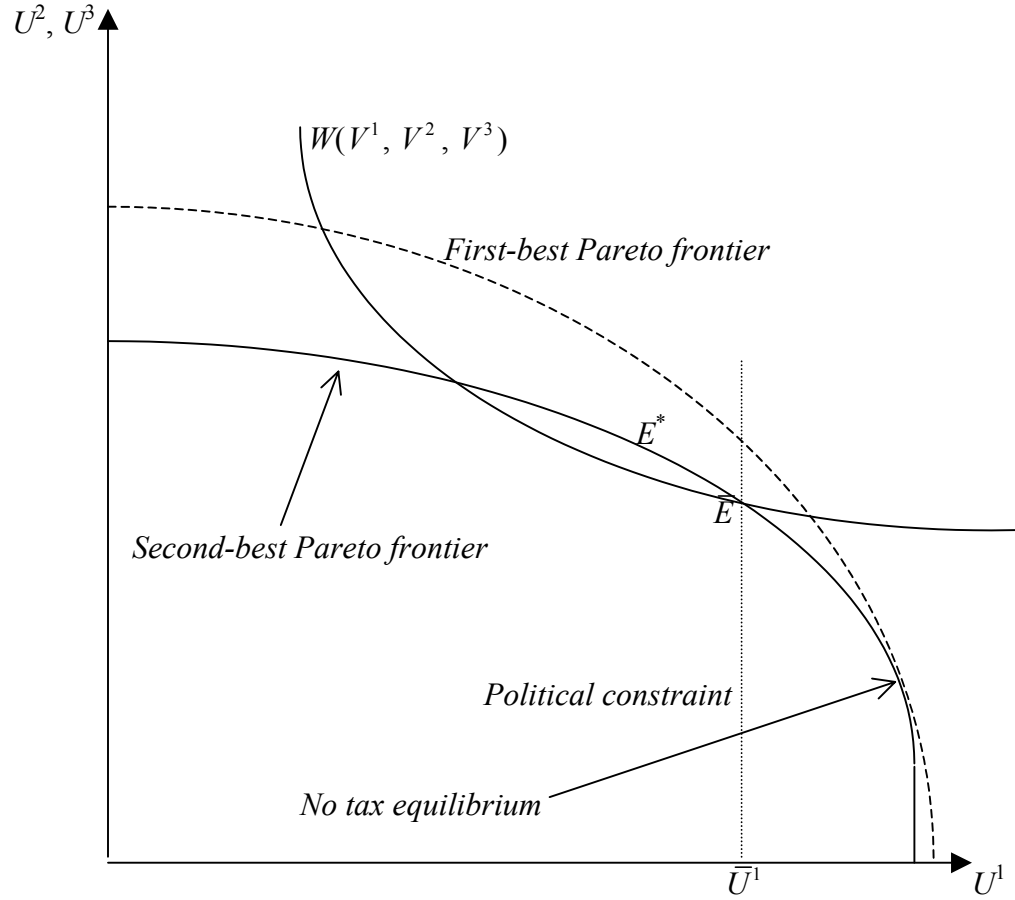


FIGURE 3

Social Welfare and Political Constraints



Appendix

(I) *Defining the Matrices and Vectors.*

The Jacobian matrix ∇Z is obtained by stacking the following matrices ∇Z^1 , ∇Z^2 , and ∇Z^3 , which correspond to the first-, second-, and third-period market clearing equations.

$$\nabla Z^1 := \begin{pmatrix} \nabla_{\pi^1} x^1 - \nabla_{\rho^1} x_1 & \nabla_{\pi^1} x^1 & \nabla_{\rho^1} x^1 - \nabla_{r^1} x_1 & \nabla_{\rho^1} x^1 & \nabla_{\omega^1} x^1 - \nabla_{w^1} x_1 & \nabla_{\omega^1} x^1 \\ \nabla_{\pi^1} k^1 - \nabla_{\rho^1} k_1 & \nabla_{\pi^1} k^1 & \nabla_{\rho^1} k^1 - \nabla_{r^1} k_1 & \nabla_{\rho^1} k^1 & \nabla_{\omega^1} k^1 - \nabla_{w^1} k_1 & \nabla_{\omega^1} k^1 \\ \nabla_{\pi^1} b^1 & \nabla_{\pi^1} b^1 & \nabla_{\rho^1} b^1 & \nabla_{\rho^1} b^1 & \nabla_{\omega^1} b^1 & \nabla_{\omega^1} b^1 \\ \nabla_{\pi^1} l^1 + \nabla_{\rho^1} L_1 & \nabla_{\pi^1} l^1 & \nabla_{\rho^1} l^1 + \nabla_{r^1} L_1 & \nabla_{\rho^1} l^1 & \nabla_{\omega^1} l^1 + \nabla_{w^1} L_1 & \nabla_{\omega^1} l^1 \end{pmatrix} \sim$$

$$\begin{pmatrix} \mathbf{0}^{(n)} & \mathbf{0}^{(n \times n)} & \mathbf{0}^{(n \times n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n \times n)} \\ 0 & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0}^{(n)} \\ -1 & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0}^{(n)} \\ 0 & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{0}^{(n)} \end{pmatrix} \sim$$

$$\begin{pmatrix} \mathbf{0}^{(n \times n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \mathbf{0}^{(n)} & \nabla_R x^1 \\ \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R k^1 \\ \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R b^1 \\ \mathbf{0}^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R l^1 \end{pmatrix}$$

$$\nabla Z^2 := \begin{pmatrix} \nabla_{\pi^1} x^2 & \nabla_{\pi^1} x^2 & \nabla_{\rho^1} x^2 & \nabla_{\rho^1} x^2 & \nabla_{\omega^1} x^2 & \nabla_{\omega^1} x^2 & \mathbf{0}^{(n)} & \nabla_{\pi^2} x^2 - \nabla_{\rho^2} x_2 \\ \nabla_{\pi^1} k^2 & \nabla_{\pi^1} k^2 & \nabla_{\rho^1} k^2 & \nabla_{\rho^1} k^2 & \nabla_{\omega^1} k^2 & \nabla_{\omega^1} k^2 & 0 & \nabla_{\pi^2} k^2 - \nabla_{\rho^2} k_2 \\ \nabla_{\pi^1} b^2 & \nabla_{\pi^1} b^2 & \nabla_{\rho^1} b^2 & \nabla_{\rho^1} b^2 & \nabla_{\omega^1} b^2 & \nabla_{\omega^1} b^2 & 0 & \nabla_{\pi^2} b^2 \\ \nabla_{\pi^1} l^2 & \nabla_{\pi^1} l^2 & \nabla_{\rho^1} l^2 & \nabla_{\rho^1} l^2 & \nabla_{\omega^1} l^2 & \nabla_{\omega^1} l^2 & 0 & \nabla_{\pi^2} l^2 + \nabla_{\rho^2} L_2 \\ -\nabla_{\pi^1} k^1 & -\nabla_{\pi^1} k^1 & -\nabla_{\rho^1} k^1 & -\nabla_{\rho^1} k^1 & -\nabla_{\omega^1} k^1 & -\nabla_{\omega^1} k^1 & 0 & \nabla_{\rho^2} K_2 \\ -\nabla_{\pi^1} b^1 & -\nabla_{\pi^1} b^1 & -\nabla_{\rho^1} b^1 & -\nabla_{\rho^1} b^1 & -\nabla_{\omega^1} b^1 & -\nabla_{\omega^1} b^1 & 0 & \mathbf{0}^{(n)} \end{pmatrix} \sim$$

$$\begin{aligned}
& \left(\begin{array}{cccccc}
\nabla_{\pi^2} x^2 & \nabla_{\rho^2} x^2 - \nabla_{r^2} x_2 & \nabla_{\rho^2} x^2 & \nabla_{\omega^2} x^2 - \nabla_{w^2} x_2 & \nabla_{\omega^2} x^2 & \nabla_{\sigma^2} x^2 - \nabla_{s^2} x_2 \\
\nabla_{\pi^2} k^2 & \nabla_{\rho^2} k^2 - \nabla_{r^2} k_2 & \nabla_{\rho^2} k^2 & \nabla_{\omega^2} k^2 - \nabla_{w^2} k_2 & \nabla_{\omega^2} k^2 & \nabla_{\sigma^2} k^2 - \nabla_{s^2} k_2 \\
\nabla_{\pi^2} b^2 & \nabla_{\rho^2} b^2 & \nabla_{\rho^2} b^2 & \nabla_{\omega^2} b^2 & \nabla_{\omega^2} b^2 & \nabla_{\sigma^2} b^2 \\
\nabla_{\pi^2} l^2 & \nabla_{\rho^2} l^2 + \nabla_{r^2} L_2 & \nabla_{\rho^2} l^2 & \nabla_{\omega^2} l^2 + \nabla_{w^2} L_2 & \nabla_{\omega^2} l^2 & \nabla_{\sigma^2} l^2 + \nabla_{s^2} L_2 \\
0^{(n)} & \nabla_{r^2} K_2 & 0 & \nabla_{w^2} K_2 & 0 & \nabla_{s^2} K_2 \\
0^{(n)} & 0 & 0 & 0 & 0 & 0
\end{array} \right) \sim \\
& \left(\begin{array}{cccccccccc}
\nabla_{\sigma^2} x^2 & 0^{(n)} & 0^{(n)} & 0^{(n \times n)} & 0^{(n \times n)} & 0^{(n)} & 0^{(n)} & 0^{(n)} & 0^{(n)} & 0^{(n)} & \nabla_R x^2 \\
\nabla_{\sigma^2} k^2 & 0 & 0 & 0^{(n)} & 0^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R k^2 \\
\nabla_{\sigma^2} b^2 & -1 & 0 & 0^{(n)} & 0^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R b^2 \\
\nabla_{\sigma^2} l^2 & 0 & 0 & 0^{(n)} & 0^{(n)} & 0 & 0 & 0 & 0 & 0 & \nabla_R l^2 \\
0 & 0 & 0 & 0^{(n)} & 0^{(n)} & 0 & 0 & 0 & 0 & 0 & -\nabla_R k^1 \\
0 & 0 & 1 & 0^{(n)} & 0^{(n)} & 0 & 0 & 0 & 0 & 0 & -\nabla_R b^1
\end{array} \right) \\
& \nabla Z^3 := \left(\begin{array}{cccccc}
\nabla_{\pi^1} x^3 & \nabla_{\pi^1} x^3 & \nabla_{\rho^1} x^3 & \nabla_{\rho^1} x^3 & \nabla_{\omega^1} x^3 & \nabla_{\omega^1} x^3 & 0^{(n)} & \nabla_{\pi^2} x^3 \\
\nabla_{\pi^1} l^3 & \nabla_{\pi^1} l^3 & \nabla_{\rho^1} l^3 & \nabla_{\rho^1} l^3 & \nabla_{\omega^1} l^3 & \nabla_{\omega^1} l^3 & 0 & \nabla_{\pi^2} l^3 \\
-\nabla_{\pi^1} k^2 & -\nabla_{\pi^1} k^2 & -\nabla_{\rho^1} k^2 & -\nabla_{\rho^1} k^2 & -\nabla_{\omega^1} k^2 & -\nabla_{\omega^1} k^2 & 0 & -\nabla_{\pi^2} k^2 \\
-\nabla_{\pi^1} b^2 & -\nabla_{\pi^1} b^2 & -\nabla_{\rho^1} b^2 & -\nabla_{\rho^1} b^2 & -\nabla_{\omega^1} b^2 & -\nabla_{\omega^1} b^2 & 0 & -\nabla_{\pi^2} b^2
\end{array} \right) \sim \\
& \left(\begin{array}{cccccccc}
\nabla_{\pi^2} x^3 & \nabla_{\rho^2} x^3 & \nabla_{\rho^2} x^3 & \nabla_{\omega^2} x^3 & \nabla_{\omega^2} x^3 & \nabla_{\sigma^2} x^3 & \nabla_{\sigma^2} x^3 & 0^{(n)} \\
\nabla_{\pi^2} l^3 & \nabla_{\rho^2} l^3 & \nabla_{\rho^2} l^3 & \nabla_{\omega^2} l^3 & \nabla_{\omega^2} l^3 & \nabla_{\sigma^2} l^3 & \nabla_{\sigma^2} l^3 & 0 \\
-\nabla_{\pi^2} k^2 & -\nabla_{\rho^2} k^2 & -\nabla_{\rho^2} k^2 & -\nabla_{\omega^2} k^2 & -\nabla_{\omega^2} k^2 & -\nabla_{\sigma^2} k^2 & -\nabla_{\sigma^2} k^2 & 0 \\
-\nabla_{\pi^2} b^2 & -\nabla_{\rho^2} b^2 & -\nabla_{\rho^2} b^2 & -\nabla_{\omega^2} b^2 & -\nabla_{\omega^2} b^2 & -\nabla_{\sigma^2} b^2 & -\nabla_{\sigma^2} b^2 & 0
\end{array} \right) \sim \\
& \left(\begin{array}{ccccc}
0^{(n)} & \nabla_{\pi^3} x^3 - \nabla_{p^3} x_3 & \nabla_{\pi^3} x^3 & \nabla_{\omega^3} x^3 - \nabla_{w^3} x_3 & \nabla_{\omega^3} x^3 \\
0 & \nabla_{\pi^3} l^3 + \nabla_{p^3} L_3 & \nabla_{\pi^3} l^3 & \nabla_{\omega^3} l^3 + \nabla_{w^3} L_3 & \nabla_{\omega^3} l^3 \\
0 & \nabla_{p^3} K_3 & 0^{(n)} & \nabla_{w^3} K_3 & 0 \\
0 & 0^{(n)} & 0^{(n)} & 0 & 0
\end{array} \right) \sim
\end{aligned}$$

$$\begin{pmatrix} \nabla_{\sigma^3} x^3 - \nabla_{s^3} x_3 & \nabla_{\sigma^3} x^3 & \mathbf{0}^{(n)} & \nabla_R x^3 \\ \nabla_{\sigma^3} l^3 + \nabla_{s^3} L_3 & \nabla_{\sigma^3} l^3 & 0 & \nabla_R l^3 \\ \nabla_{s^3} K_3 & 0 & 0 & -\nabla_R k^2 \\ 0 & 0 & 1 & -\nabla_R b^2 \end{pmatrix}$$

The gradients of the consumers' indirect utility functions (with respect to dP) are given by:

$$\nabla V^1 := \langle \nabla_{\pi^1} V^1, \nabla_{\pi^1} V^1, \nabla_{\rho^1} V^1, \nabla_{\rho^1} V^1, \nabla_{\omega^1} V^1, \nabla_{\omega^1} V^1, \mathbf{0}^{(4n+14)}, \nabla_R V^1 \rangle$$

$$\nabla V^2 := \langle \nabla_{\pi^1} V^2, \nabla_{\pi^1} V^2, \nabla_{\rho^1} V^2, \nabla_{\rho^1} V^2, \nabla_{\omega^1} V^2, \nabla_{\omega^1} V^2, \mathbf{0}, \nabla_{\pi^2} V^2, \nabla_{\pi^2} V^2, \nabla_{\rho^2} V^2, \nabla_{\rho^2} V^2, \nabla_{\omega^2} V^2, \nabla_{\omega^2} V^2, \nabla_{\sigma^2} V^2, \nabla_{\sigma^2} V^2, \mathbf{0}^{(2n+7)}, \nabla_R V^2 \rangle$$

$$\nabla V^3 := \langle \nabla_{\pi^1} V^3, \nabla_{\pi^1} V^3, \nabla_{\rho^1} V^3, \nabla_{\rho^1} V^3, \nabla_{\omega^1} V^3, \nabla_{\omega^1} V^3, \mathbf{0}, \nabla_{\pi^2} V^3, \nabla_{\pi^2} V^3, \nabla_{\rho^2} V^3, \nabla_{\rho^2} V^3, \nabla_{\omega^2} V^3, \nabla_{\omega^2} V^3, \nabla_{\sigma^2} V^3, \nabla_{\sigma^2} V^3, \mathbf{0}^{(2)}, \nabla_{\pi^3} V^3, \nabla_{\pi^3} V^3, \nabla_{\omega^3} V^3, \nabla_{\omega^3} V^3, \nabla_{\sigma^3} V^3, \nabla_{\sigma^3} V^3, \mathbf{0}, \nabla_R V^3 \rangle$$

The gradient of the social welfare function (with respect to dP) is given by:

$$\begin{aligned} \nabla W := & \langle \lambda_1 \nabla_{\pi^1} V^1 + \lambda_2 \nabla_{\pi^1} V^2 + \lambda_3 \nabla_{\pi^1} V^3, \lambda_1 \nabla_{\pi^1} V^1 + \lambda_2 \nabla_{\pi^1} V^2 + \lambda_3 \nabla_{\pi^1} V^3, \lambda_1 \nabla_{\rho^1} V^1 \\ & + \lambda_2 \nabla_{\rho^1} V^2 + \lambda_3 \nabla_{\rho^1} V^3, \lambda_1 \nabla_{\rho^1} V^1 + \lambda_2 \nabla_{\rho^1} V^2 + \lambda_3 \nabla_{\rho^1} V^3, \lambda_1 \nabla_{\omega^1} V^1 + \lambda_2 \nabla_{\omega^1} V^2 + \lambda_3 \nabla_{\omega^1} V^3, \\ & \lambda_1 \nabla_{\omega^1} V^1 + \lambda_2 \nabla_{\omega^1} V^2 + \lambda_3 \nabla_{\omega^1} V^3, \mathbf{0}, \lambda_2 \nabla_{\pi^2} V^2 + \lambda_3 \nabla_{\pi^2} V^3, \lambda_2 \nabla_{\pi^2} V^2 + \lambda_3 \nabla_{\pi^2} V^3, \lambda_2 \nabla_{\rho^2} V^2 + \\ & \lambda_3 \nabla_{\rho^2} V^3, \lambda_2 \nabla_{\rho^2} V^2 + \lambda_3 \nabla_{\rho^2} V^3, \lambda_2 \nabla_{\omega^2} V^2 + \lambda_3 \nabla_{\omega^2} V^3, \lambda_2 \nabla_{\omega^2} V^2 + \lambda_3 \nabla_{\omega^2} V^3, \lambda_2 \nabla_{\sigma^2} V^2 + \\ & \lambda_3 \nabla_{\sigma^2} V^3, \lambda_2 \nabla_{\sigma^2} V^2 + \lambda_3 \nabla_{\sigma^2} V^3, \mathbf{0}^{(2)}, \lambda_3 \nabla_{\pi^3} V^3, \lambda_3 \nabla_{\pi^3} V^3, \lambda_3 \nabla_{\omega^3} V^3, \lambda_3 \nabla_{\omega^3} V^3, \lambda_3 \nabla_{\sigma^3} V^3, \\ & \lambda_3 \nabla_{\sigma^3} V^3, \mathbf{0}, \lambda_1 \nabla_R V^1 + \lambda_2 \nabla_R V^2 + \lambda_3 \nabla_R V^3 \rangle \end{aligned}$$

where:

$$\lambda_i = \frac{\partial W(\cdot)}{\partial V^i}$$

is the welfare weight of generation i .

(II) Proof of Corollary 2.1.

Subtracting the first equation in (4.1) from the second yields:

$$\langle \mu_1, \dots, \mu_n \rangle \nabla_{p^1} x_1 + \mu_{n+1} \nabla_{p^1} k_1 - \mu_{n+3} \nabla_{p^1} L_1 = 0^{(n)} \quad (\text{A.1})$$

Subtracting the third equation in (4.1) from the fourth yields:

$$\langle \mu_1, \dots, \mu_n \rangle \nabla_{r^1} x_1 + \mu_{n+1} \nabla_{r^1} k_1 - \mu_{n+3} \nabla_{r^1} L_1 = 0 \quad (\text{A.2})$$

Subtracting the fifth equation in (4.1) from the sixth yields:

$$\langle \mu_1, \dots, \mu_n \rangle \nabla_{w^1} x_1 + \mu_{n+1} \nabla_{w^1} k_1 - \mu_{n+3} \nabla_{w^1} L_1 = 0 \quad (\text{A.3})$$

Combining (A.1) – (A.3) yields:

$$\langle \mu_1, \dots, \mu_n, \mu_{n+1}, \mu_{n+3} \rangle \nabla_{\hat{p}^1} \hat{x}_1 = 0^{(n+2)} \quad (\text{A.4})$$

where \hat{x}_1 combines the supply of x_1 , k_1 , and L_1 , and \hat{p}^1 combines the producer prices p^1 , r^1 , and w^1 .

Next, it follows from homogeneity of degree zero of $x_1(p^1, r^1, w^1)$, $k_1(p^1, r^1, w^1)$, and $L_1(p^1, r^1, w^1)$ in producer prices that:

$$\nabla_{\hat{p}^1} \hat{x}_1 \langle p_1^1, \dots, p_n^1, r^1, w^1 \rangle = 0^{(n+2)} \quad (\text{A.5})$$

It now follows from (A.4), (A.5), and symmetry of the Hessian of the profit function $\psi_1(p^1, r^1, w^1)$ that $\langle \mu_1, \dots, \mu_n, \mu_{n+1}, \mu_{n+3} \rangle = \eta \langle p_1^1, \dots, p_n^1, r^1, w^1 \rangle$ for some scalar $\eta > 0$.

That is, first-period social shadow prices are proportional to first-period producer prices.

In a similar manner, it can be shown that second- and third-period social shadow prices are proportional to second- and third-period producer prices. ■

(III) Proof of Proposition 3.

The proof is analogous to that of Corollary 2.1, except subtracting the third equation in (5.1) from the fourth yields:

$$\langle \mu_1, \dots, \mu_n \rangle \nabla_{r^1} x_1 + \mu_{n+1} \nabla_{r^1} k_1 - \mu_{n+3} \nabla_{r^1} L_1 = \theta_1 \quad (\text{A.2}')$$

The expression corresponding to (A.4) becomes:

$$\langle \mu_1, \dots, \mu_n, \mu_{n+1}, \mu_{n+3} \rangle \nabla_{\hat{p}^1} \hat{x}_1 = \langle 0^{(n)}, \theta_1, 0 \rangle \quad (\text{A.4}')$$

From (A.4') and (A.5) it follows that first-period social shadow prices are proportional to first-period producer prices if and only if $\theta_1 = 0$, i.e., if and only if the constraint preventing increases in first-period capital taxation is not binding. In a similar manner, it can be shown that production efficiency holds in the second period if and only if $\theta_2 = 0$, i.e., if and only if the constraint preventing increases in second-period capital taxation is not binding. ■

(IV) Proof of Proposition 4.

Requiring that the government balance its budget period-by-period requires that the following restrictions be placed on its bond issues: $db_1 + b_1 = 0$, $db_2 + b_2 = 0$, $dB_2 + B_2 = 0$, and $dB_3 + B_3 = 0$. There exists an equilibrium preserving and Pareto improving reform with the government's budget balanced period-by-period if and only if there exists a solution dP to the following non-homogenous system:

$$\nabla Z dP \leq 0^{(3n+11)} \quad \nabla V dP \gg 0^{(3)} \quad \begin{pmatrix} 0^{(2n+4)} & 1 & 0^{(4n+14)} \\ 0^{(4n+11)} & 1 & 0^{(2n+7)} \\ 0^{(4n+12)} & 1 & 0^{(2n+6)} \\ 0^{(6n+17)} & 1 & 0 \end{pmatrix} dP = \begin{pmatrix} -b_1 \\ -b_2 \\ -B_2 \\ -B_3 \end{pmatrix} \quad (\text{A.6})$$

The non-homogenous system (A.6) has a solution dP if and only if the following homogeneous system (A.7) has a solution $\langle dP, D \rangle$, where D is a dummy variable used to convert the non-homogenous system (A.6) into the homogeneous system (A.7).

$$\begin{pmatrix} \nabla Z & 0^{(3n+11)} \end{pmatrix} \begin{pmatrix} dP \\ D \end{pmatrix} \leq 0^{(3n+11)} \quad \begin{pmatrix} \nabla V & 0^{(3)} \\ 0^{(6n+19)} & 1 \end{pmatrix} \begin{pmatrix} dP \\ D \end{pmatrix} \gg 0^{(4)}$$

$$\begin{pmatrix} 0^{(2n+4)} & 1 & 0^{(4n+14)} & b_1 \\ 0^{(4n+11)} & 1 & 0^{(2n+7)} & b_2 \\ 0^{(4n+12)} & 1 & 0^{(2n+6)} & B_2 \\ 0^{(6n+17)} & 1 & 0 & B_3 \end{pmatrix} \begin{pmatrix} dP \\ D \end{pmatrix} = 0^{(4)} \quad (\text{A.7})$$

By Motzkin's Theorem, if there is no solution $\langle dP, D \rangle$ to (A.7), then there exist vectors $\langle \lambda, \xi \rangle > 0^{(4)}$, $\mu \geq 0^{(3n+11)}$, and θ unrestricted such that:

$$\begin{pmatrix} \lambda & \xi \end{pmatrix} \begin{pmatrix} \nabla V & 0^{(3)} \\ 0^{(6n+19)} & 1 \end{pmatrix} + \theta \begin{pmatrix} 0^{(2n+4)} & 1 & 0^{(4n+14)} & b_1 \\ 0^{(4n+11)} & 1 & 0^{(2n+7)} & b_2 \\ 0^{(4n+12)} & 1 & 0^{(2n+6)} & B_2 \\ 0^{(6n+17)} & 1 & 0 & B_3 \end{pmatrix} = \mu \begin{pmatrix} \nabla Z & 0^{(3n+11)} \end{pmatrix} \quad (\text{A.8})$$

The only differences between (A.8) and the unrestricted conditions for second-best optimality contained in (4.1) are that θ_1 , θ_2 , θ_3 , and θ_4 are added, respectively, to the 7th, 16th, 17th, and 24th equations in (4.1), and an additional condition:

$$\xi + \theta_1 b_1 + \theta_2 b_2 + \theta_3 B_2 + \theta_4 B_3 = 0$$

is introduced. However, none of these changes affect the equations required to prove production efficiency (see the proof of Corollary 2.1). Therefore, restrictions on the government's debt instruments do not affect the desirability of production efficiency. ■

References

- Ahmad, E. and N. Stern [1984], "The Theory of Reform and Indian Indirect Taxes", *Journal of Public Economics*, 25, 259-298.
- Blackorby, C. and C. Brett [1999], "The Pareto Optima of Finite-Horizon OLG Models", *University of British Columbia Discussion Paper*, 99-27.
- Blackorby, C. and C. Brett [2000], "Fiscal Federalism Revisited", *Journal of Economic Theory*, 92, 300-317.
- Blackorby, C. and C. Brett [2004], "Capital Taxation in a Simple Finite-Horizon OLG Model", *Warwick Economic Research Papers*, 709.
- Blanchard, O. and S. Fischer [1989], *Lectures on Macroeconomics*, MIT Press.
- Brett, C [1998], "Tax Reform and Collective Family Decision-Making", *Journal of Public Economics*, 70, 425-440.
- Diamond, P. and J. Mirrlees [1971], "Optimal Taxation and Public Production", *American Economic Review*, 61(1), 8-27 and 61(3), 261-278.
- Diewert, W [1978], "Optimal Tax Perturbations", *Journal of Public Economics*, 10, 139-177.
- Dixit, A [1979], "Price Changes and Optimum Taxation in a Many-Consumer Economy", *Journal of Public Economics*, 11, 143-157.
- Fogelman, F., M. Quinzii, and R. Guesnerie [1978], "Dynamic Processes for Tax Reform Theory", *Journal of Economic Theory*, 17, 200-226.
- Guesnerie, R [1977], "On the Direction of Tax Reform", *Journal of Public Economics*, 7, 179-202.

- Guesnerie, R [1995], *A Contribution to the Pure Theory of Taxation*, Cambridge University Press.
- Hu, S [1979], “Social Security, the Supply of Labor, and Capital Accumulation”, *American Economic Review*, 69, 274-283.
- Keen, M. and D. Wildasin [2004], “Pareto-Efficient International Taxation”, *American Economic Review*, 94(1), 259-275.
- King, M [1983], “Welfare Analysis of Tax Reforms using Household Data”, *Journal of Public Economics*, 21, 183-214.
- Lipsey, R. and K. Lancaster [1956], “The General Theory of Second Best”, *Review of Economic Studies*, 24, 11-32.
- Mangasarian, O [1969], *Nonlinear Programming*, McGraw-Hill.
- Murty, S. and R. Russell [2003], “Externality Policy Reform: A General Equilibrium Analysis”, *Journal of Public Economic Theory*, forthcoming.
- Weymark, J [1979], “A Reconciliation of Recent Results in Optimal Taxation”, *Journal of Public Economics*, 7, 171-190.