Using distance functions to derive optimal progressive earnings tax and commodity tax structures

John Burbidge*

11 December 2018

Abstract

Much of the research program in optimal taxation rests on the Atkinson-Stiglitz theorem (1976) — in the presence of optimal nonlinear earnings taxation, if leisure is weakly separable from goods, there is no role for differential commodity taxation. The nonlinear earnings tax in the theorem is one where, conditional on reported earnings, the government can choose tax paid and the marginal tax rate (mtr). The relationship between the average tax rate (atr) and mtr is unrestricted. Most governments operate *progressive* nonlinear tax systems in which, for each person paying taxes, mtr is not less than atr. I build on Deaton's work on distance functions and taxation to show that the AS theorem fails in the presence of optimal progressive earnings taxation. Conditional on mtr $\geq$ atr, the search for optimal earnings tax structures cannot be undertaken without simultaneously studying optimal commodity taxation whether or not leisure is weakly separable from goods. The formal theory in the paper assumes two types. I also discuss a finite-type example of an optimal progressive earnings, and commodity, tax structure and present numerical examples with four types.

Keywords: *Optimal taxation, distance function, separability*

JEL claasification H21

*Department of Economics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, jburbidg@uwaterloo.ca. I thank Lutz-Alexander Busch, Christoph Lülfesmann, Steeve Mongrain, John Revesz, César Sosa-Padilla, Michael Veall, and seminar participants at McMaster University, the University of Waterloo, the 2014 and 2016 Canadian Economics Association Meetings and the 2017 Insitut d'Economia de Barcelona Workshop on the Economics of Taxation for helpful comments.

1 Introduction

Decades ago public economists used to believe that a government's ability to vary proportional tax rates across commodities was a prominent feature of an efficient revenue structure. In the most recent Handbook of Public Economics Piketty and Saez explain how the Atkinson-Stiglitz theorem changed this.

'Atkinson and Stiglitz (1976) derived the very important and influential result that under separability and homogeneity assumptions on preferences, differentiated commodity taxation is not useful when earnings can be taxed nonlinearly. This famous result was influential both for shaping the field of optimal tax theory and in tax policy debates. Theoretically, it contributed greatly to shift the theoretical focus toward optimal nonlinear taxation and away from the earlier Diamond and Mirrlees (1971) model of differentiated commodity taxation (itself based on the original Ramsey (1927) contribution). Practically, it gave a strong rationale for eliminating preferential taxation of necessities on redistributive grounds, and using instead a uniform value-added-tax combined with income-based transfers and *progressive income taxation.* Even more importantly, the Atkinson and Stiglitz (1976) result has been used to argue against the taxation of capital income and in favor of taxing solely earnings or consumption.' (Piketty and Saez (2013), pp. 401-402; my emphasis)[1]

Given the Atkinson-Stiglitz theorem, and its assumption that leisure is weakly separable from goods, public economists have set to one side the question of optimal commodity taxation to focus on determining optimal *progressive* earnings tax structure. But in a *progressive* earnings tax system, for those making net contributions to the tax-transfer system, the average earnings tax rate, atr, rises with earnings and the marginal earnings tax rate, mtr, must exceed the atr. I build on Deaton's work on distance functions and taxation to show that the Atkinson-Stiglitz theorem requires the possibility that, for those making net contributions to the tax-transfer system, atr *exceed* mtr. It turns out that even if leisure is additively separable from goods, determining the optimal *progressive* earnings tax structure cannot be separated from determining the optimal commodity tax structure.

I begin with the Ramsey (1927) tax problem which is about choosing proportional tax rates when some items in the consumer's budget constraint, e.g. leisure, cannot be taxed directly. Assume the simplest setting possible: two goods and leisure. The government might tax either good or earnings and, given the individual's budget constraint, one must choose some normalization of tax rates: there is no loss in generality in setting the earnings tax rate to zero to focus on optimal commodity taxation. Absent an instrument to tax leisure directly the government must try to tax leisure indirectly by setting a higher tax rate on whichever good is more complementary with leisure. It is well known that the deadweight loss of tax rates arises solely from substitution effects so the "complementarity" between goods and leisure must be measured with utility held constant. The distance

[1]There is a literature on the optimality of capital taxation; e.g. Erosa and Gervais (2002), Conesa, Kitao and Krueger (2009) and Burbidge (2015).

function is ideally suited to this task. In this context it is the number by which a bundle of goods and leisure must be scaled to deliver a particular level of utility, say u_0; the arguments of the distance function are the same as those of the ordinary utility function, plus u_0. The observation that distance functions offer the clearest understanding of optimal commodity taxes was made by Deaton in a series of papers that includes Deaton (1979, 1981). Among other things, he derived optimal commodity-tax formulae that showed goods more complementary with leisure should be taxed at higher rates and he proved the formulae were equivalent to the better-known inverse elasticity rule.

A distance-function approach to solving the Ramsey problem can be implemented by choosing goods and leisure to maximize government revenue subject to certain constraints, one of which must be a lower bound on utility. If one modified the Ramsey problem to allow an unrestricted nonlinear earnings tax the only constraint would be the lower bound on utility. The math would dictate setting all commodity tax rates and the mtr to zero and collecting all revenue with a lump-sum tax on earnings; atr positive, mtr zero. If one were to constrain the nonlinear earnings tax to be progressive then one would have to ensure that the mtr be at least as large as the atr. Since this constraint would bind the optimal earnings tax would be proportional (mtr = atr). Then, following the logic of the previous paragraph, the proportional earnings tax rate can be set to zero and the math dictates setting the commodity tax rate on good 1 higher than the tax rate good 2 if good 1 is more complementary with leisure than is good 2. In section 2, I show that constraining the earnings tax to be progressive — mtr $\geq$ atr — is straightforward with distance functions.

The Mirrlees (1971) problem is about taxation for redistribution with imperfect information — the government can observe earnings but not wage rates or hours worked. Assume the simplest setting possible — two types A and B, with $w^A > w^B$. In Model 1 in this paper, the government chooses goods and leisure for each type to maximize total revenue subject to minimum-utility constraints for each A and for each B (u_0^A, u_0^B), a constraint that As and Bs pay the same prices for goods, and constraints that yield the outcome that the As not mimic the Bs. The value function for this problem gives maximum revenue as a function of u_0^A, u_0^B and other parameters. Setting this function to zero implicitly defines u_0^B as a function of u_0^A, that is, the utility possibility frontier, upf.[2] I show that Model 1 delivers the well-known results in the literature, which include a zero marginal earnings tax rate on the high earner, and the Atkinson-Stiglitz theorem — zero proportional tax rates on goods 1 and 2 if leisure is weakly separable from goods. Even in Model 1, however, there is a strong link to the Ramsey problem. In the Ramsey problem,

[2]Every point on the upf will be associated with a particular bundle of goods and leisure for each A and another bundle of goods and leisure for each B. And for each point on the upf there will be a continuum of equivalent tax systems depending on whatever normalization of tax rates is chosen. A distinct advantage of the distance-function approach over other approaches is that its instruments *are* the variables that determine each point on the upf. I *derive* optimal tax results in terms of relations between marginal rates of substitution and ratios of private equilibrium prices — p_1, p_2, w^A, w^B, which are assumed constant. And then, for each model, I *present* the results using the normalization that makes the results easiest to understand for those familiar with the optimal tax literature.

and, if leisure is not weakly separable from goods in Model 1, good 1 should be taxed at a higher rate than good 2 if, holding utility constant, good 1 is more complementary with leisure (see Jacobs and Boadway (2014)).

Notice that Model 1 does not restrict the nonlinear earnings tax to be progressive — for the high earner atr exceeds mtr. On the upf, between the private equilibrium and the point where the no-mimicking constraints bind, Model 1 delivers first-best results: lump-sum taxes on the high earners fund a cash transfer to each low earner. After the mimicking constraints bind it is efficient to supplement lump-sum taxes on the high earners with a marginal earnings tax rate on the low earners to discourage them from working, together with the possibility of different proportional tax rates on goods 1 and 2 if leisure is not weakly separable from goods.

Model 2 in this paper is identical to Model 1 except that I impose the constraint that the earnings tax system be progressive — mtr $\geq$ atr for the high earner. The addition of a binding constraint forces the upf for this model inside, or at best to touch, the upf of Model 1 (it will touch at the private equilibrium). If one adopts the normalization that the mtr for low earners is zero, I prove that, along the upf, from the private equilibrium to the point where mimicking starts, mtr = atr for high earners and this earnings tax on high earners is the primary source of revenue for cash payments to the low earners. Even if leisure is weakly separable from goods (but preferences are not homothetic)[3] the tax rate will be positive on whichever good is more complementary with leisure, and negative on the other good. Along the upf, after the mimicking constraints bind, one can show that the earnings tax rate on the high earners is reduced to discourage mimicking, and increases in goods' tax rates gradually replace the earnings tax rate on the high earners as the source of revenue for redistribution. The Ramsey rule holds along the entire upf — the commodity tax rate is higher on whichever good is more complementary with leisure.

Does Model 2 generalize to three or more types? With three or more types there is, of course, a continuum of paths along the upf. I work out the optimal earnings and commodity tax structure for a particular path along the upf with a finite number of types, and present numerical examples with four types. This particular path exhibits the main results of Model 2.

In the next section I model the Ramsey problem in the presence of nonlinear earnings taxes. Section 3 uses the framework of section 2 to study the Mirrlees problem with Models 1, 2 and the extension of 2 to a finite number of types. Section 4 summarizes and concludes.

2 Ramsey with a nonlinear earnings tax

Denote goods consumption and leisure by (x_1, x_2, l), prices without taxes by (p_1, p_2, w), the time endowment by L, proportional goods' tax rates by (t_1, t_2), and the nonlinear earnings

[3] In the appendix I prove that in both models in this paper weak separability between leisure and goods together with homothetic preferences imply equal tax rates on goods 1 and 2.

tax function by $T(w(L-l))$. The individual's budget constraint can be written as

$$(1+t_1)\, p_1 x_1 + (1+t_2)\, p_2 x_2 \quad = \quad w(L-l) - T(w(L-l)). \tag{1}$$

If $u(x_1, x_2, l)$ is the ordinary utility function the distance function, $d(x_1, x_2, l, u_0)$, is defined by

$$u\left(\frac{x_1}{d(x_1,x_2,l,u_0)}, \frac{x_2}{d(x_1,x_2,l,u_0)}, \frac{l}{d(x_1,x_2,l,u_0)}\right) = u_0, \tag{2}$$

that is, $d(x_1, x_2, l, u_0)$ is the number by which any vector of goods and leisure must be scaled to deliver utility level u_0, and

$$u(x_1, x_2, l) \geq u_0 \text{ if and only if } d(x_1, x_2, l, u_0) \geq 1. \tag{3}$$

Deaton (1979) showed the distance function and the expenditure function, $e\,(p_1, p_2, w, u_0)$, are related to each other in the following way.

$$e\,(p_1, p_2, w, u_0) = \underset{x_1, x_2, l}{\text{Min}} \; \frac{p_1 x_1 + p_2 x_2 + wl}{d(x_1, x_2, l, u_0)}. \tag{4}$$

$$d\,(x_1, x_2, l, u_0) = \underset{p_1, p_2, w}{\text{Min}} \; \frac{p_1 x_1 + p_2 x_2 + wl}{e\,(p_1, p_2, w, u_0)}. \tag{5}$$

The derivatives of the expenditure function with respect to prices are Hicksian demands, $h_i\,(p_1, p_2, w, u_0)\,, i = 1, 2, 3$, which map from prices to quantities. The derivatives of the distance function with respect to quantities, denote these by $a_i\,(x_1, x_2, l, u_0)\,, i = 1, 2, 3$, are like inverse Hicksian demands that map from quantities to prices. Applying the envelope theorem to (5) with tax rates in place obtain

$$a_1\,(x_1, x_2, l, u_0) \quad \equiv \quad \frac{\partial d\,(x_1, x_2, l, u_0)}{\partial x_1} = \frac{(1+t_1)p_1}{E} \tag{6}$$

$$a_2\,(x_1, x_2, l, u_0) \quad \equiv \quad \frac{\partial d\,(x_1, x_2, l, u_0)}{\partial x_2} = \frac{(1+t_2)p_2}{E} \tag{7}$$

$$a_3\,(x_1, x_2, l, u_0) \quad \equiv \quad \frac{\partial d\,(x_1, x_2, l, u_0)}{\partial l} = \frac{(1-T')\,w}{E}. \tag{8}$$

In this context E is

$$(1+t_1)\, p_1 x_1 + (1+t_2)\, p_2 x_2 + \big(1 - T'\,(w(L-l))\big)\, wl = E. \tag{9}$$

Deaton proved $a_i/a_j, i \neq j$ is the marginal rate of substitution between i and j, denoted MRS_{ij}. Just as the Hessian of the expenditure function, the Slutsky matrix, must be symmetric and negative semi-definite, so must the Hessian of the distance function, the *Antonelli* matrix, be symmetric and negative semi-definite. In addition, pre-multiplying the Slutsky matrix by prices yields a vector of zeros and pre-multiplying the Antonelli matrix by $[x_1 \; x_2 \; l]$ yields a row vector of zeros.[4]

The Ramsey problem can be solved by maximizing government revenue, $wL - p_1 x_1 - p_2 x_2 - wl$, subject to constraints. If the only constraint is a lower bound on utility[5] — utility $\geq u_0$ or $d(x_1, x_2, l, u_0) \geq 1$ — the math will deliver an answer equivalent to setting $t_1 = t_2 = T' = 0$ and collecting all revenue, $T(w(L-l))$, with a lump-sum tax. Here the atr is positive and the mtr is zero. What happens if we constrain the nonlinear earnings tax to be progressive, mtr $\geq$ atr?

Drop the argument of the earnings tax function and its derivative. Then rewrite (9) as

$$(1+t_1)\, p_1 x_1 + (1+t_2)\, p_2 x_2 = E - \left(1 - T'\right) wl$$

and from the budget constraint

$$(1+t_1)\, p_1 x_1 + (1+t_2)\, p_2 x_2 = w\,(L-l) - T.$$

Thus

$$E = \left(1 - T'\right) wl + w\,(L-l) - T.$$

If the atr equals the mtr then $T = T'w(L-l)$, $E = (1-T')wL$ and using (8), $a_3 = 1/L$. If, as in a progressive nonlinear earnings tax system, mtr $\geq$ atr,

$$\frac{1}{L} - a_3 \geq 0, \tag{10}$$

and the Lagrange multiplier on this constraint must be positive. The Lagrangian for the government's problem can be written as

$$\mathcal{L} \;=\; wL - p_1 x_1 - p_2 x_2 - wl + \lambda^u \left(d\left(x_1, x_2, l, u_0\right) - 1\right) + \lambda^L \left(\frac{1}{L} - a_3\left(x_1, x_2, l, u_0\right)\right).$$

The first-order conditions are

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x_1} = -p_1 + \lambda^u a_1 - \lambda^L a_{31} \\ 0 &= \frac{\partial \mathcal{L}}{\partial x_2} = -p_2 + \lambda^u a_2 - \lambda^L a_{32} \\ 0 &= \frac{\partial \mathcal{L}}{\partial l} = -w + \lambda^u a_3 - \lambda^L a_{33} \end{aligned}$$

[4]Following Deaton (1981), I assume the elements on the main diagonal of the Antonelli matrix are negative and the individual consumes a positive amount of each good (pages 1248 and 1251).

[5]Assume u_0 is less than utility in the laisser-faire equilibrium.

Using the symmetry of the Antonelli matrix

$$\lambda^u = \frac{p_1}{a_1} + \lambda^L a_{13}^* = \frac{p_2}{a_2} + \lambda^L a_{23}^* = \frac{w}{a_3} + \lambda^L a_{33}^*,$$

where $a_{ij}^* \equiv a_{ij}/a_i$. From the second equality

$$\frac{p_2}{a_1}\left(\frac{a_1}{a_2} - \frac{p_1}{p_2}\right) \quad = \quad \lambda^L \left(a_{13}^* - a_{23}^*\right). \tag{11}$$

The sign of $a_1/a_2 - p_1/p_2$ matches the sign of $a_{13}^* - a_{23}^*$. a_{j3}^* measures the degree of complementarity between good j and leisure *holding utility constant*, and from (6) and (7), $a_1/a_2 = (1+t_1)p_1/((1+t_2)p_2)$. The algebra confirms the Corlett-Hague (1953-54) intuition that good 1 should be taxed at a higher rate than good 2 if good 1 is more complementary with leisure than is good 2, and vice versa. The algebra also shows that the optimal progressive nonlinear earnings tax should be proportional with mtr equal to atr.

It is worth noting that weak separability between goods and leisure does not imply $a_{13}^* = a_{23}^*$. For example, Auerbach (1979) showed, for any positive prices, $t_1 > t_2$ ($a_{13}^* > a_{23}^*$) is always optimal with

$$u(x_1, x_2, l) = (x_1 x_2)^{1/2} + x_1^{1/2} + l^{1/2}.$$

3 Taxation for redistribution

Consider an economy with two types of price-taking agents like the agent discussed above. As and Bs differ only in their wage rates with $w^A > w^B$. Earnings have to be spent on goods 1 and 2, which are taxed at proportional rates t_1, t_2. Assume the government wishes to redistribute money from the As to the Bs but it can observe only an individual's earnings, not the individual's wage rate. This is the Mirrlees (1971) problem with two types. To derive the characteristics of efficient allocations suppose the government maximizes government revenue subject to minimum-utility constraints for A and B and various other constraints listed below. The value function for this problem gives maximized revenue as a function of the A and B utility levels and other parameters such as the time endowment. If the only purpose of taxation is to raise revenue for redistribution setting this value function to zero implicitly defines B's utility as a function of A's utility, that is, the upf.

The government's ability to redistribute efficiently depends on the instruments at its disposal. The equivalent of equations (6)-(8) in the present context are

$$a_1^A = \frac{(1+t_1)\,p_1}{E^A};\; a_2^A = \frac{(1+t_2)\,p_2}{E^A};\; a_3^A = \frac{\left(1-t^H\right)w^A}{E^A} \tag{12}$$

$$a_1^B = \frac{(1+t_1)\,p_1}{E^B};\; a_2^B = \frac{(1+t_2)\,p_2}{E^B};\; a_3^B = \frac{\left(1-t^L\right)w^B}{E^B}, \tag{13}$$

where E^j is the total "expenditure" of agent $j = A, B$ and $t^j, j = H, L$ are mtrs for high and low earners.[6]

Government revenue is

$$n^A\left(w^A L - p_1 x_1^A - p_2 x_2^A - w^A l^A\right) + n^B\left(w^B L - p_1 x_1^B - p_2 x_2^B - w^B l^B\right),$$

$n^j, j = A, B$ is the number of each type. Let $\lambda^j, j = A, B$ be the Lagrange multipliers on the minimum utility constraints for A and B:[7]

$$d\left(x_1^j, x_2^j, l^j, u_0^j\right) - 1 \quad \geq \quad 0.\; j = A, B \tag{14}$$

Each low earner will be given a cash transfer equal to total revenue divided by n^B. $T^L \leq 0$ denotes total earnings "tax" paid by each low earner.

As noted above I assume the government cannot condition the proportional commodity tax rates on person type — the As and Bs pay the same prices for the two goods. From (12) and (13)

$$\frac{(1+t_1)\,p_1}{(1+t_2)\,p_2} = \frac{a_1^A\left(x_1^A, x_2^A, l^A, u_0^A\right)}{a_2^A\left(x_1^A, x_2^A, l^A, u_0^A\right)} = \frac{a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right)}{a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right)}$$

or

$$\begin{aligned} &a_1^A\left(x_1^A, x_2^A, l^A, u_0^A\right) a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) - \\ &a_2^A\left(x_1^A, x_2^A, l^A, u_0^A\right) a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right) = 0. \end{aligned} \tag{15}$$

Denote the Lagrange multiplier on this constraint by λ^p. At this point I cannot place restrictions on the sign of this multiplier but more can be said in the models presented later in the paper.

I turn now to mimicking constraints. Starting at the private equilibrium where all tax rates are zero there is no incentive for an A to mimic a B. But as tax rates are increased and the revenue given to the low earners one moves up the utility possibility frontier (draw the upf with u^A on the horizontal axis) in the direction of lower A utility. At some point each A will realize that utility would be higher if the A pretended to be a low earner and

[6] Note that these equalities build in the assumption that A and B pay the same prices for goods.

[7] These must be positive; see footnote 13.

was eligible for the cash transfer. When an A mimics a B, the mimicking A chooses leisure, l^{AB}, to make earnings $w^A(L - l^{AB})$ equal the earnings of a B, $w^B(L - l^B)$. By mimicking a B the mimicking A receives the cash transfer, $-T^L$, and will face a budget constraint for goods 1 and 2 that is identical to that faced by each B. With mimicking there are three extra constraints and two new choice variables, x_1^{AB} and x_2^{AB}. One of the extra constraints is that A's utility acting as an A be at least as high as the utility of an A mimicking a B, that is, $u\left(x_1^A, x_2^A, l^A\right) \geq u\left(x_1^{AB}, x_2^{AB}, l^{AB}\right)$. This constraint together with the constraint that $u\left(x_1^A, x_2^A, l^A\right) \geq u_0^A$ can be written as (14) and (16).

$$1 - d\left(x_1^{AB}, x_2^{AB}, l^{AB}, u_0^A\right) \geq 0 \tag{16}$$

The Lagrange multiplier on (16), λ^{AB}, must be nonnegative.

A second constraint is that mimicking As pay the same prices for goods as everyone else. Then the other side of the observation that to prevent mimicking the government must make the utility of an A at least as large as the utility of an A mimicking a B is that, to discourage mimicking, the government would like to have an instrument that would discourage mimicking by pushing the goods budget of a mimicking A below the goods budget for a B. Using (13) and $a_1^{AB}/a_2^{AB} = a_1^B/a_2^B$ this could be written as

$$\begin{aligned} &a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_1^B + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_2^B > \\ &a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_1^{AB} + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_2^{AB}. \end{aligned}$$

The *absence* of such an instrument means that the second and third constraints arising from mimicking can be collapsed into

$$\begin{aligned} &a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_1^B + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_2^B \leq \\ &a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_1^{AB} + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) x_2^{AB} \text{ or} \\ &a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right)\left(x_1^B - x_1^{AB}\right) + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right)\left(x_2^B - x_2^{AB}\right) \leq 0. \end{aligned} \tag{17}$$

The Lagrange multiplier on this constraint, λ^{ABc}, must be non-positive.

Before going into the details of the two models it is useful to think about goods and leisure choices on the upf above the point where mimicking constraints bind (recall that I assumed u^A is on horizontal axis). Consider a step up the upf. u^B must increase and the fall in u^A must equal the fall in u^{AB}. If dl^B were negative or zero, from $w^A(L - l^{AB}) = w^B(L - l^B)$, it follows that $dl^{AB} = (w^B/w^A)dl^B$, that is, l^{AB} would have to fall by less or the same amount as l^B. Therefore, for u^B to rise, the goods budget for B would have to increase but this goods budget is the same for a mimicking A and a B, and the marginal utility of leisure is lower for a mimicking A than a B. There may be utility functions where, moving up the upf after mimicking begins, l^B decreases (in particular, leisure would have to be an inferior good) but I am going to rule them out. I assume that l^B and l^{AB} increase on the upf after the point where mimicking begins and the reason u^{AB} falls as u^B increases, even though at each step the goods' budgets change by the same amount, is that $dl^{AB} = (w^B/w^A)dl^B < dl^B$. As in the previous section, many of the results in this section depend on the complementarity of goods and leisure. I assume the sign of $a^{*j}_{13} - a^{*j}_{23}$ is the same for $j = A, B, AB$.

3.1 Model 1

In Model 1 the government uses x^A_1, x^A_2, l^A, x^B_1, x^B_2, l^B to maximize government revenue subject to (14) and (15) if mimicking constraints do not bind. If they do the government has two extra choice variables, x^{AB}_1, x^{AB}_2, to deal with the three mimicking constraints, (16), (17) and $a^{AB}_1/a^{AB}_2 = a^B_1/a^B_2$.[8] Since the optimal earnings tax literature has focused on a zero marginal earnings tax rate for the high earner a 'natural' normalization to choose for this model is $t^H = 0$.

Statement 1: (a) Along the upf, from the private equilibrium until the mimicking constraints bind, $T^H \geq 0, t^L = t_1 = t_2 = 0$. (b) After mimicking constraints bind: (i) $T^H > 0, t^L > 0$; (ii) if leisure and goods are weakly separable $t_1 = t_2 = 0$ (the Atkinson-Stiglitz theorem); (iii) if leisure and goods are not weakly separable t_1 and $-t_2$ have the sign of $a^{*j}_{13} - a^{*j}_{23}, j = A, B$.[9]

In this model, between the private equilibrium and the point where mimicking constraints bind, the upf is first best — a lump-sum tax on high earners, $T^H > 0$, funds a cash transfer to low earners, $T^L < 0$. After mimicking constraints bind we move onto a second-best upf inside the first-best upf. T^H is still positive and T^L is negative but now other taxes help deal with the mimicking constraints. First, the marginal earnings tax rate on the low earner is positive, $t^L > 0$. This tax induces the low earners to work less which reduces the incentive for high earners to mimic low earners. Second, with T^H available and leisure weakly separable from goods, there is no role for differential commodity taxation — the Atkinson-Stiglitz theorem. But if leisure is not weakly separable from goods *and* $a^{*j}_{13} \neq a^{*j}_{23}$ then commodity taxation helps. For example, if good 1 is more complementary with leisure than good 2, $a^{*j}_{13} > a^{*j}_{23}$, then taxing good 1 and subsidizing good 2 penalizes mimicking As who consume more leisure than As. This particular result is consistent with the findings of Jacobs and Boadway (2014).

Readers of earlier drafts of this paper have asked whether the optimality of uniform commodity taxation in Model 1 implies leisure must be weakly separable from goods. The answer is no. Consider an extension of the example in Auerbach (1979).

[8] Recall that l^{AB} is chosen to make the earnings of a mimicking A equal the earnings of a B.
[9] The proof is in the Appendix.

$$u(x_1, x_2, l) = (x_1 x_2)^{1/2} + (x_1 l)^{1/2} + (x_2 l)^{1/2},$$

when $p_1 = p_2$. The perfect symmetry between each good and leisure implies the optimality of uniform commodity taxation and $a_{13}^{*j} = a_{23}^{*j}$ and yet leisure is not weakly separable from goods.

3.2 Model 2

This model is the same as Model 1 except, for the high earner, I impose the restriction that the mtr equals or exceeds the atr. The extra constraint which is the equivalent of (10) is

$$\frac{1}{L} - a_3^A \left(x_1^A, x_2^A, l^A, u_0^A\right) \geq 0. \tag{18}$$

And, of course, one must choose a normalization for tax rates. The one that makes the results most transparent is to set the marginal earnings tax rates on the low earner, t^L, equal to zero. Statement 2 lists the implications of Model 2.

Statement 2: (a) Whether or not leisure is weakly separable from goods, if $a_{13}^{*j} = a_{23}^{*j}$: (i) t^H rises from the private equilibrium until the mimicking constraints bind, and thereafter it falls; (ii) $t_1 = t_2 = 0$ until the mimicking constraints bind and thereafter $t_1 = t_2$ rise; (b) if $a_{13}^{*j} > a_{23}^{*j}$ the results in (a) are modified only in that $t_1 > 0$ and $t_2 < 0$ before the mimicking constraints bind, and $t_1 > t_2$ after the mimicking constraints bind; the opposite is true if $a_{13}^{*j} < a_{23}^{*j}$.

Model 2 imposes the restriction that for the high earner the mtr exceed or equal atr. Given the results of Model 1 this constraint binds. Starting at the private equilibrium along the upf we are in a second-best world where the primary instrument for raising revenue for redistribution is a proportional earnings tax on the high earner. This tax of course discourages work effort of the high earner and so it helps to supplement $t^H > 0$ with commodity taxation of leisure, e.g. $t_1 > 0$ and $t_2 < 0$ if good 1 is more complementary with leisure, $a_{13}^{*j} > a_{23}^{*j}$. After mimicking constraints bind we are in a third-best world where it's most efficient to raise further redistribution revenue using commodity taxation and to reduce t^H to deal with the mimicking constraints.

3.3 Model 2 with a finite number of types

Model 2 solves the Mirrlees problem for progressive earnings taxation and two types. If Model 2 could not be generalized to a finite number of types it would hold limited interest. With three or more types the upf is a surface with a continuum of paths leading away from the private equilibrium. This section examines a particular path along a upf with a finite number of types and presents numerical results for a four-type example. In Model 2 with two types I did not have to say much about the progressive earnings tax other than to specify that, for each type, mtr could not be less than atr. In this section I assume this is true for each type and in addition, looking across types, the mtr cannot decrease when moving to a higher wage-rate type, which is a standard feature of a progressive earnings tax system. As in Model 2 I normalize tax rates by setting the mtr for the lowest-wage type to zero.

I begin with the simplest case where $a_{13}^{*j} = a_{23}^{*j}$ because then $t_1 = t_2$ along any path on the upf. At the private equilibrium the equivalent of (18) holds with equality for all types. At the first step the utility of the highest-wage type is reduced and that of the lowest-wage type is increased; the utility of all other types is held constant. To accomplish this the government uses a proportional earnings tax rate on the type with the highest wage rate, label this t^A for a type A, to fund a cash transfer to the lowest earner.[10] Inspection of the equivalent of (13) for the lowest earner reveals that a_3 falls because the numerator is constant and expenditure in the denominator increases; $a_3 = 1/L$ for all other types.

Proceeding in this direction mimicking constraints will eventually bind for the highest two earners or the lowest two earners. Suppose it is the latter. Efficient ways to deal with these constraints are to reduce the mtr for the second highest earner or increase the mtr for the lowest earner. But neither is feasible in a progressive earnings tax system where mtrs cannot fall as earnings rise. So just enough cash must be given to the second lowest earner to prevent this person from mimicking the lowest earner. Now $1/L > a_3$ for the second lowest earner $> a_3$ for the lowest earner and $a_3 = 1/L$ for everyone else.

Further redistribution will eventually cause mimicking constraints to bind either for the second and third lowest earners or for the two highest earners. Suppose it is the latter. At this point $t^A > t^B = 0$. As in Model 1 the most efficient way to cope with these mimicking constraints is to induce Bs to work less, take more leisure — increase t^B and give Bs some cash. Now $1/L = a_3^A > a_3^B$. So further redistribution continues with t^B rising until it equals t^A; under a progressive earnings tax t^B cannot exceed t^A. Once this constraint binds further redistribution requires that mtrs on the highest two earners increase lock step. The particular path along the upf described here holds B utility constant until the mtrs for A and B move together.

[10]Everywhere else in this paper I use a notation that tries to emphasize the point that the government cannot identify a person's type, thus t^H rather than t^A etc. But in this section it is convenient to label mtrs $t^j, j = A, B, \ldots$.

To this point, as in Model 2, $t_1 = t_2 = 0$. And, again as in Model 2, when redistribution proceeds to the point where mimicking constraints bind for those being taxed and those being subsidized, further redistribution occurs with earnings tax rates being reduced to cope with mimicking constraints and the lost revenue is made up by increasing commodity tax rates, still with $t_1 = t_2$. Table 1 presents an example of this path along a four-person upf with Cobb-Douglas utility, where $a_{13}^{*j} = a_{23}^{*j}$.[11]

When $a_{13}^{*j} \neq a_{23}^{*j}$ the story is modified slightly. In the first stage, if $a_{13}^{*j} > a_{23}^{*j}$, then it is efficient that $t_1 > t_2$ to tax leisure to counteract increases in t^A which encourage the As to work less. As in Model 2, where it is also true that the mtr of the lowest earner is zero, this is implemented with $t_1 > 0, t_2 < 0$. The optimal levels of these tax rates trade off the benefit of getting the As to work harder against the distortions to labour supply and consumption plans for everyone else. At the first step away from the private equilibrium the utility of As falls, the utility of lowest earners rises, and everyone else's utility is held constant. The most efficient way to hold utility constant for types in between the lowest and highest wage earners is to increase the earnings tax rates slightly to offset the commodity-tax-rate effects on leisure and to pay each earnings level some cash. For example, see line 2 of Tables 2 and 3 both of which have utility functions where $a_{13}^{*j} > a_{23}^{*j}$. This process implies that when mimicking constraints bind for the two lowest earners the second lowest earners have a positive mtr. At this point, the most efficient way to deal with these mimicking constraints as further redistribution is carried out is to reduce the second-lowest earners' mtr. Once it hits zero the next step of redistribution proceeds as described above with mtrs equal to zero for the two lowest earning groups. This extra step is line 3 in Tables 2 and 3 and it has no counterpart in Table 1, where $a_{13}^{*j} = a_{23}^{*j}$.

In section 2 I showed that for any particular type $\text{mtr}^i \geq \text{atr}^i$ is equivalent to $1/L - a_3^i \geq 0$. How does one use distance functions to capture the constraints that $t^i \geq t^j$ if $w^i > w^j$? From the extensions of (12) and (13)

$$(1 + t_1)\, p_1 = \frac{\left(1 - t^i\right) w^i a_1^i}{a_3^i} = \frac{\left(1 - t^j\right) w^j a_1^j}{a_3^j}$$

Thus

$$t^i \geq t^j \text{ if and only if } w^i a_1^i a_3^j - w^j a_1^j a_3^i \geq 0.$$

[11] If $u(x_1, x_2, l) = x_1^{\alpha} x_2^{\beta} l^{1-\alpha-\beta}, \alpha > 0, \beta > 0, \alpha + \beta < 1$ then $d(x_1, x_2, l, u_0) = u_0^{-1} x_1^{\alpha} x_2^{\beta} l^{1-\alpha-\beta}$ and $a_{13}^{*j} = (1 - \alpha - \beta)/l^j = a_{23}^{*j}$.

4 Summary and conclusions

This paper uses a simple framework to show that the Atkinson-Stiglitz theorem does not hold when those making net contributions to the tax-transfer program face optimal progressive earnings taxation. The decade of the 1970s was a long time ago and much has changed. To take one example, a large literature has been built on Diamond (1980) and Saez (2002) who introduced search and other frictions into the competitive labour markets assumed in Mirrlees (1971). Nevertheless, it is difficult to exaggerate the influence of the Atkinson-Stiglitz theorem on the modern research program in public economics. Even though the governments in these models are very instrument-constrained, researchers typically do not consider the role that optimal commodity taxation might play in improving social outcomes. The main purpose of this paper is to persuade researchers in public economics to reconsider their decision to ignore commodity tax instruments.

Just as the Ramsey problem does not vanish in the Mirrlees framework, I suspect the tension between arranging commodity taxes to tax leisure, or to tax work, holding utility constant, will be present in models with various frictions in labour or other markets. It is likely that distance functions will continue to be a useful analytical device — recall that distance functions are a transformation of the utility function, and therefore can be used whether agents buy and sell in competitive markets or are constrained in the transactions they can make. Furthermore, Deaton (1981) describes precisely how to estimate the parameters necessary to implement tax rules based on distance functions. The combination of better analytical tools that can be implemented empirically should yield new insights into tax policy.

Appendix

Model 1 is a special case of Model 2. The optimization problem for Model 2 with mimicking constraints is

$$\underset{\substack{x_1^A, x_2^A, l^A \\ x_1^B, x_2^B, l^B \\ x_1^{AB}, x_2^{AB} \\ \lambda^j, j=A,B,p,L, \\ AB, ABc}}{\textbf{Opt}} \quad n^A\left(w^A L - p_1 x_1^A - p_2 x_2^A - w^A l^A\right) + n^B\left(w^B L - p_1 x_1^B - p_2 x_2^B - w^B l^B\right) +$$
$$\lambda^A\left(d\left(x_1^A, x_2^A, l^A, u_0^A\right) - 1\right) +$$
$$\lambda^B\left(d\left(x_1^B, x_2^B, l^B, u_0^B\right) - 1\right) +$$
$$\lambda^p\left(a_1^A\left(x_1^A, x_2^A, l^A, u_0^A\right) a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right) - a_2^A\left(x_1^A, x_2^A, l^A, u_0^A\right) a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right)\right) +$$
$$\lambda^L\left(\frac{1}{L} - a_3^A\left(x_1^A, x_2^A, l^A, u_0^A\right)\right) +$$
$$\lambda^{AB}\left(1 - d\left(x_1^{AB}, x_2^{AB}, l^{AB}, u_0^A\right)\right) +$$
$$\lambda^{ABc}\left(a_1^B\left(x_1^B, x_2^B, l^B, u_0^B\right)\left(x_1^B - x_1^{AB}\right) + a_2^B\left(x_1^B, x_2^B, l^B, u_0^B\right)\left(x_2^B - x_2^{AB}\right)\right)$$

where

$$w^A\left(L - l^{AB}\right) \quad = \quad w^B\left(L - l^B\right) \tag{19}$$

The first-order conditions for the eight goods and leisure variables are

$$\begin{aligned}
0 &= -n^A p_1 + \lambda^A a_1^A + \lambda^p\left(a_2^B a_{11}^A - a_1^B a_{21}^A\right) - \lambda^L a_{31}^A \\
0 &= -n^A p_2 + \lambda^A a_2^A + \lambda^p\left(a_2^B a_{12}^A - a_1^B a_{22}^A\right) - \lambda^L a_{32}^A \\
0 &= -n^A w^A + \lambda^A a_3^A + \lambda^p\left(a_2^B a_{13}^A - a_1^B a_{23}^A\right) - \lambda^L a_{33}^A
\end{aligned}$$

$$
\begin{aligned}
0 &= -n^B p_1 + \lambda^B a_1^B + \lambda^p \left(a_1^A a_{21}^B - a_2^A a_{11}^B\right) + \\
&\quad \lambda^{ABc}\left(a_{11}^B\left(x_1^B - x_1^{AB}\right) + a_{21}^B\left(x_2^B - x_2^{AB}\right) + a_1^B\right) \\
0 &= -n^B p_2 + \lambda^B a_2^B + \lambda^p \left(a_1^A a_{22}^B - a_2^A a_{12}^B\right) + \\
&\quad \lambda^{ABc}\left(a_{12}^B\left(x_1^B - x_1^{AB}\right) + a_{22}^B\left(x_2^B - x_2^{AB}\right) + a_2^B\right) \\
0 &= -n^B w^B + \lambda^B a_3^B + \lambda^p \left(a_1^A a_{23}^B - a_2^A a_{13}^B\right) + \\
&\quad -\lambda^{AB} a_3^{AB}\frac{w^B}{w^A} + \lambda^{ABc}\left(a_{13}^B\left(x_1^B - x_1^{AB}\right) + a_{23}^B\left(x_2^B - x_2^{AB}\right)\right) \\
0 &= -\lambda^{AB} a_1^{AB} - \lambda^{ABc} a_1^B \\
0 &= -\lambda^{AB} a_2^{AB} - \lambda^{ABc} a_2^B.
\end{aligned}
$$

Note the last two equations imply that[12]

$$
\begin{aligned}
\lambda^{ABc} a_1^B &= -\lambda^{AB} a_1^{AB} \\
\lambda^{ABc} a_2^B &= -\lambda^{AB} a_2^{AB} \\
\lambda^{ABc} &= -\lambda^{AB}\frac{E^B}{E^{AB}}.
\end{aligned}
$$

Using these equations obtain

$$
\begin{aligned}
n^A p_1 &= \lambda^A a_1^A + \lambda^p \left(a_2^B a_{11}^A - a_1^B a_{21}^A\right) - \lambda^L a_{31}^A && (20) \\
n^A p_2 &= \lambda^A a_2^A + \lambda^p \left(a_2^B a_{12}^A - a_1^B a_{22}^A\right) - \lambda^L a_{32}^A && (21) \\
n^A w^A &= \lambda^A a_3^A + \lambda^p \left(a_2^B a_{13}^A - a_1^B a_{23}^A\right) - \lambda^L a_{33}^A && (22) \\
n^B p_1 &= \lambda^B a_1^B + \lambda^p \left(a_1^A a_{21}^B - a_2^A a_{11}^B\right) - \\
&\quad \lambda^{AB}\frac{E^B}{E^{AB}}\left(a_{11}^B\left(x_1^B - x_1^{AB}\right) + a_{21}^B\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} a_1^{AB} && (23) \\
n^B p_2 &= \lambda^B a_2^B + \lambda^p \left(a_1^A a_{22}^B - a_2^A a_{12}^B\right) - \\
&\quad \lambda^{AB}\frac{E^B}{E^{AB}}\left(a_{12}^B\left(x_1^B - x_1^{AB}\right) + a_{22}^B\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} a_2^{AB} && (24) \\
n^B w^B &= \lambda^B a_3^B + \lambda^p \left(a_1^A a_{23}^B - a_2^A a_{13}^B\right) - \\
&\quad \lambda^{AB}\frac{E^B}{E^{AB}}\left(a_{13}^B\left(x_1^B - x_1^{AB}\right) + a_{23}^B\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} a_3^{AB}\frac{w^B}{w^A}. && (25)
\end{aligned}
$$

Recall $a_{ij}^* \equiv a_{ij}/a_i$ and use (12), (13) and the symmetry of the Antonelli matrix to obtain

[12] E^{AB} is the cost of buying $(x_1^{AB}, x_2^{AB}, l^{AB})$ at $((1+t_1)p_1, (1+t_2)p_2, w^{AB})$. w^{AB} is the shadow price of leisure for a mimicking A and equals $(1+t_1)p_1 a_3^{AB}/a_1^{AB}$.

$$n^A \frac{p_1}{a_1^A} = n^A E^A \frac{1}{1+t_1} = \lambda^A + \lambda^p \left(a_2^B a_{11}^{*A} - a_1^B a_{12}^{*A}\right) - \lambda^L a_{13}^{*A} \tag{26}$$

$$n^A \frac{p_2}{a_2^A} = n^A E^A \frac{1}{1+t_2} = \lambda^A + \lambda^p \left(a_2^B a_{21}^{*A} - a_1^B a_{22}^{*A}\right) - \lambda^L a_{23}^{*A} \tag{27}$$

$$n^A \frac{w^A}{a_3^A} = \lambda^A + \lambda^p \left(a_2^B a_{31}^{*A} - a_1^B a_{32}^{*A}\right) - \lambda^L a_{33}^{*A} \tag{28}$$

$$\begin{aligned} n^B \frac{p_1}{a_1^B} &= n^B E^B \frac{1}{1+t_1} = \lambda^B + \lambda^p \left(a_1^A a_{12}^{*B} - a_2^A a_{11}^{*B}\right) - \\ \lambda^{AB} \frac{E^B}{E^{AB}} &\left(a_{11}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{12}^{*B}\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} \frac{E^B}{E^{AB}} \end{aligned} \tag{29}$$

$$\begin{aligned} n^B \frac{p_2}{a_2^B} &= n^B E^B \frac{1}{1+t_2} = \lambda^B + \lambda^p \left(a_1^A a_{22}^{*B} - a_2^A a_{21}^{*B}\right) - \\ \lambda^{AB} \frac{E^B}{E^{AB}} &\left(a_{21}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{22}^{*B}\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} \frac{E^B}{E^{AB}} \end{aligned} \tag{30}$$

$$\begin{aligned} n^B \frac{w^B}{a_3^B} &= \lambda^B + \lambda^p \left(a_1^A a_{32}^{*B} - a_2^A a_{31}^{*B}\right) - \\ \lambda^{AB} \frac{E^B}{E^{AB}} &\left(a_{31}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{32}^{*B}\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB} \frac{a_3^{AB}}{a_3^B} \frac{w^B}{w^A}. \end{aligned} \tag{31}$$

Then (27) minus (26), and (30) minus (29) yield[13]

$$\begin{aligned} n^A E^A \frac{t_1 - t_2}{(1+t_1)(1+t_2)} &= \lambda^p \left(a_2^B\left(a_{21}^{*A} - a_{11}^{*A}\right) + a_1^B\left(a_{12}^{*A} - a_{22}^{*A}\right)\right) + \\ \lambda^L \left(a_{13}^{*A} - a_{23}^{*A}\right) & \end{aligned} \tag{32}$$

$$\begin{aligned} n^B E^B \frac{t_1 - t_2}{(1+t_1)(1+t_2)} &= \lambda^p \left(a_1^A\left(a_{22}^{*B} - a_{12}^{*B}\right) + a_2^A\left(a_{11}^{*B} - a_{21}^{*B}\right)\right) + \\ \lambda^{AB} \frac{E^B}{E^{AB}} &\left(\left(a_{11}^{*B} - a_{21}^{*B}\right)\left(x_1^B - x_1^{AB}\right) + \left(a_{12}^{*B} - a_{22}^{*B}\right)\left(x_2^B - x_2^{AB}\right)\right). \end{aligned} \tag{33}$$

Finally, subtract $n^A E^A$ times (33) from $n^B E^B$ times (32), divide by E^B, and use (12) and (13) to write

[13] At the private equilibrium (in any of the models) all Lagrange multipliers are zero except λ^A and λ^B, and $E^j = w^j L, j = A, B$. Using (26) and (29), $\lambda^j = n^j w^j L, j = A, B$.

$$
\begin{aligned}
0 = {} & n^B \lambda^L a_1^A \left(a_{13}^{*A} - a_{23}^{*A}\right) + \\
& \lambda^p \left\{ n^A \left[a_1^B \left(a_{12}^{*B} - a_{22}^{*B}\right) + a_2^B \left(a_{21}^{*B} - a_{11}^{*B}\right)\right] + n^B \left[a_1^B \left(a_{12}^{*A} - a_{22}^{*A}\right) + a_2^B \left(a_{21}^{*A} - a_{11}^{*A}\right)\right]\right\} + \\
& \lambda^{AB} \frac{n^A a_1^{AB}}{a_1^A} \left[\left(a_{21}^{*B} - a_{11}^{*B}\right)\left(x_1^B - x_1^{AB}\right) + \left(a_{22}^{*B} - a_{12}^{*B}\right)\left(x_2^B - x_2^{AB}\right)\right]. \qquad (34)
\end{aligned}
$$

I begin with some observations that simplify the proofs of Statements 1 and 2.

R1: the coefficient of λ^p is positive in (32) and (34) and negative in (33)

In footnote (4) I noted that I follow Deaton (1981) in assuming that there are no corner solutions and that the main diagonal of the Antonelli matrix is negative. These assumptions, together with the result that multiplying any Antonelli matrix by the consumption vector yields a vector of zeros, implies that a 'typical' off-diagonal element is positive. Inspection of the preceding equations reveals terms that fit the pattern

$$a_{ij}^* - a_{jj}^* = \frac{a_{ij}}{a_i} - \frac{a_{jj}}{a_j},\ i \neq j.$$

In what follows I presume these terms are positive, and given this assumption, **R1** follows.

R2: If leisure is weakly separable from goods then

$$x_i^{AB} = x_i^B \ i = 1, 2.$$

A mimicking A and a B have the same budget for goods 1 and 2; they differ only because the mimicking A chooses more leisure. If leisure is weakly separable from goods the condition that

$$\mathrm{MRS}_{12} = \frac{(1+t_1)\,p_1}{(1+t_2)\,p_2}$$

is the same equation for both and thus the result must hold.

R3: Recall that I have assumed the sign of $a_{13}^{*j} - a_{23}^{*j}$ is independent of type $j = A, B, AB$. If leisure is not weakly separable from goods then

$$\text{Sign } a_{13}^{*j} - a_{23}^{*j} = \text{Sign } x_1^{AB} - x_1^B = \text{Sign } -\left(x_2^{AB} - x_2^B\right) = -\text{Sign } a_{31}^{*j}\left(x_1^B - x_1^{AB}\right) + a_{32}^{*j}\left(x_2^B - x_2^{AB}\right).$$

Deaton (1981) showed that the $\mathrm{MRS}_{12} = a_1/a_2$. Extending the argument in $R2$ if

$$\frac{\partial}{\partial l}\left(\frac{a_1^j}{a_2^j}\right)=\frac{a_{13}^j a_2^j - a_1^j a_{23}^j}{\left(a_2^j\right)^2}=\frac{a_1^j}{a_2^j}\left(a_{13}^{*j}-a_{23}^{*j}\right)=0,$$

then goods 1 and 2 are equally complementary with leisure and a mimicking A and a B will choose the same levels for goods 1 and 2. If $a_{13}^{*j}-a_{23}^{*j}>0$ then the mimicking A who enjoys more leisure than the B will choose more of the good most complementary with leisure, which in this case is good 1, and less of the other good. To prove the last equality in **R3** note

$$\begin{aligned}
&a_{31}^{*j}\left(x_1^B-x_1^{AB}\right)+a_{32}^{*j}\left(x_2^B-x_2^{AB}\right)=\\
&\frac{1}{a_3^j}\left(a_{13}^j\left(x_1^B-x_1^{AB}\right)+a_{23}^j\left(x_2^B-x_2^{AB}\right)\right)=\\
&\frac{1}{a_3^j}\left(a_{13}^{*j}a_1^j\left(x_1^B-x_1^{AB}\right)+a_{23}^{*j}a_2^j\left(x_2^B-x_2^{AB}\right)\right).
\end{aligned}$$

If $a_{13}^{*j}=a_{23}^{*j}$ then the last line equals

$$\frac{a_{13}^{*j}}{a_3^j}\left(a_1^j\left(x_1^B-x_1^{AB}\right)+a_2^j\left(x_2^B-x_2^{AB}\right)\right)=0,$$

because a mimicking A and a B have the same budget constraint for goods. Using the first part of this result if $a_{13}^{*j}\gtrless a_{23}^{*j}$ then $x_1^{AB}\gtrless x_1^B$ and the expression immediately above is $\lesseqgtr 0$.

R4: the λ^p term in (28) has the same sign as $a_{13}^{*A}-a_{23}^{*A}$ and the λ^p term in (31) has the same sign as $a_{23}^{*B}-a_{13}^{*B}$.
The λ^p term in (28) is

$$\begin{aligned}
a_2^B a_{31}^{*A}-a_1^B a_{32}^{*A} &= a_2^B\frac{a_{31}^A}{a_3^A}-a_1^B\frac{a_{32}^A}{a_3^A}\\
&=\frac{a_1^B}{a_3^A}\left(\frac{a_2^B}{a_1^B}a_{31}^A-a_{32}^A\right)\\
&=\frac{a_1^B}{a_3^A}\left(\frac{a_2^A}{a_1^A}a_{13}^A-a_{23}^A\right) \quad \text{A matrix symmetric; A and B pay the same prices for goods}\\
&=\frac{a_1^B a_2^A}{a_3^A}\left(\frac{a_{13}^A}{a_1^A}-\frac{a_{23}^A}{a_2^A}\right)\\
&=\frac{a_1^B a_2^A}{a_3^A}\left(a_{13}^{*A}-a_{23}^{*A}\right)
\end{aligned}$$

The proof for the λ^p term in (31) is similar.

R5: If preferences are homothetic and leisure is weakly separable from goods then in either Model 1 or 2 uniform taxation of goods is efficient.
If preferences are homothetic then there is no loss in generality in assuming the utility function, $u(x_1, x_2, l)$, is homogeneous of degree 1.

$$u(\alpha x_1, \alpha x_2, \alpha l) = \alpha u(x_1, x_2, l), \text{ for all } \alpha > 0.$$

Let α equal the inverse of the distance function. Then from the definition of the distance function

$$u\left(\frac{x_1}{d(x_1, x_2, l, u_0)}, \frac{x_2}{d(x_1, x_2, l, u_0)}, \frac{l}{d(x_1, x_2, l, u_0)}\right) = u_0$$

we have

$$d(x_1, x_2, l, u_0) = u_0^{-1} u(x_1, x_2, l) \equiv u_0^{-1} f(x_1, x_2)\, g(l).$$

Then

$$a_{13}^{*q} = \frac{g'(l^q)}{g(l^q)} = a_{23}^{*q}, \ q = A, B.$$

From here **R1**, **R2** and (34) imply $\lambda^p = 0$, and then (32) implies $t_1 = t_2$.

Proof of Statement 1

In this model λ^L is zero. Given the normalization that $t^H = 0$ equation (12) implies that $w^A = a_3^A E^A$. On the upf between the private equilibrium and the point where mimicking starts, $\lambda^{AB} = 0$. **R1** says the coefficient of λ^p in (34) is positive so (34) implies that $\lambda^p = 0$. Given the normalization, (28) implies $\lambda^A = n^A E^A$ and then $t_1 = t_2 = 0$ follows from (26) and (27). From here (29) implies $\lambda^B = n^B E^B$ and then (13) and (31) imply $t^L = 0$. This completes the proof of 1(a).

What happens along the upf after the point where mimicking begins — $\lambda^{AB} > 0$? If leisure is weakly separable from goods then using **R2** $x_i^{AB} = x_i^B, i = 1, 2$, and then (34) tells us λ^p must be zero and $t_1 = t_2 = 0$ follows, as above. This proves Statement 1(b)(ii), which is the Atkinson-Stiglitz theorem.

R6: when $\lambda^{AB} > 0$, $t_1 = t_2 = 0$ is efficient if and only if $x_i^{AB} = x_i^B, i = 1, 2$.

If $x_i^B = x_i^{AB}, i = 1, 2$ (34) implies $\lambda^p = 0$ and then $t_1 = t_2 = 0$ follows from (26) to (28). Going in the other direction, if $t_1 = t_2$ is optimal, (32) implies $\lambda^p = 0$. Then (26) to (28) imply $a_i^A / a_3^A = p_i / w^A$ and (33) implies

$$\left(a_{11}^{*B}-a_{21}^{*B}\right)\left(x_1^B-x_1^{AB}\right)+\left(a_{12}^{*B}-a_{22}^{*B}\right)\left(x_2^B-x_2^{AB}\right)=0.$$

Using **R1**, if $(x_1^B - x_1^{AB})$ were positive (negative) then $(x_2^B - x_2^{AB})$ would be negative (positive) and then above equation could not be true. Therefore, $x_i^{AB} = x_i^B, i = 1, 2$.

Now combine **R1**, **R3** and **R6** in the context of Model 1. If leisure is not weakly separable from goods and $\lambda^{AB} > 0$, making $a_{13}^{*j} - a_{23}^{*j}, j = A, B$ positive in turn makes the coefficient of λ^{AB} in (34) negative and therefore λ^p positive. Using **R1** again, (26) and (27) imply $t_1 > 0$ and $t_2 < 0$. Making $a_{13}^{*j} - a_{23}^{*j}, j = A, B$ negative moves each tax rate in the opposite direction. This proves Statement 1 (b) (iii).

I now move on to the proof that when $\lambda^{AB} > 0$, $t^L > 0$. Use (13), (29), (30), (31) and $(1+t_1)p_1 = a_1^B E^B = a_1^{AB} E^{AB}$ to write

$$\frac{p_i}{a_i^B}\frac{a_3^B}{w^B} = \frac{1-t^L}{1+t_i}$$

$$= \frac{\lambda^B + \lambda^p\left(a_1^A a_{i2}^{*B} - a_2^A a_{i1}^{*B}\right) - \lambda^{AB}\frac{E^B}{E^{AB}}\left(a_{i1}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{i2}^{*B}\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB}\frac{a_1^{AB}}{a_1^B}}{\lambda^B + \lambda^p\left(a_1^A a_{32}^{*B} - a_2^A a_{31}^{*B}\right) - \lambda^{AB}\frac{E^B}{E^{AB}}\left(a_{31}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{32}^{*B}\left(x_2^B - x_2^{AB}\right)\right) - \lambda^{AB}\frac{a_3^{AB}}{a_3^B}\frac{w^B}{w^A}}, i = 1, 2 \tag{35}$$

Case 1: Leisure is weakly separable from goods, or leisure is not weakly separable from goods and $a_{13}^{*j} = a_{23}^{*j}$. These assumptions imply $t_i = 0$ and the middle two terms in the numerator and denominator on the RHS of (35) are zero. Then at the point on the upf where the mimicking constraints bind λ^{AB} becomes positive. The LHS of the equation above falls below 1 ($t^L > 0$) if and only if

$$\frac{a_1^{AB}}{a_1^B} > \frac{a_3^{AB}}{a_3^B}\frac{w^B}{w^A}. \tag{36}$$

This inequality holds. a_1/a_3 is the MRS between good 1 and leisure, and mimicking As enjoy more leisure and the same utility as As, so

$$\frac{a_1^{AB}}{a_3^{AB}} > \frac{a_1^A}{a_3^A} \text{ or } \frac{a_1^{AB}}{a_3^{AB}}\frac{a_3^B}{a_1^B}\frac{w^A}{w^B} > \frac{a_1^A}{a_3^A}\frac{a_3^B}{a_1^B}\frac{w^A}{w^B}.$$

and, using (12) and (13), the right side equals unity when mimicking starts.

Case 2: Leisure is not weakly separable from goods and $a_{13}^{*j} \neq a_{23}^{*j}, j = A, B$. Use (13) to rewrite (31):

$$\left(n^B + \lambda^{AB}\frac{a_3^{AB}}{w^A}\right)\frac{1}{1-t^L} = \frac{\lambda^B}{E^B} + \frac{\lambda^p}{E^B}\left(a_1^A a_{32}^{*B} - a_2^A a_{31}^{*B}\right) - \frac{\lambda^{AB}}{E^{AB}}\left(a_{31}^{*B}\left(x_1^B - x_1^{AB}\right) + a_{32}^{*B}\left(x_2^B - x_2^{AB}\right)\right).$$

Take the case where we increase $a_{13}^{*j} - a_{23}^{*j}$ from zero to a small positive number. In (34) the coefficient of λ^p and λ^{AB} are positive numbers. Using **R1**, **R3** and **R4**, in (34), the coefficient of λ^{AB} is a small negative number and therefore λ^p must be a small positive number. Using **R4** the entire λ^p term in the equation above will be negative but it will be second-order of smalls (two small numbers multiplied together). The entire λ^{AB} term will be a small positive number. Thus the effect of making $a_{13}^{*j} - a_{23}^{*j}$ positive is to increase the RHS of the equation which means t^L becomes more positive. This completes the proof of Statement 1 (i).

Proof of Statement 2

In this model all Lagrange multipliers may be nonzero and recall that the chosen normalization of tax rates is $t^L = 0$.

Case 1, $a_{13}^{*j} = a_{23}^{*j}$: Using **R2** and **R3**, whether or not leisure is weakly separable from goods, $x_j^{AB} = x_j^B, j = 1, 2$. Then (34) implies $\lambda^p = 0$. Thus the middle two terms of the numerator and denominator of (35) are zero. Before mimicking constraints bind (35) implies $t_1 = t_2 = 0$. After mimicking constraints bind the argument in case 1 of the proof of Statement 1 implies $t_1 = t_2$ rise as λ^{AB} rises from zero. What happens to t^H along the upf?

The ratio of (28) to (26) or (27) in the present context is

$$\frac{1+t_j}{1-t^H} = \frac{\lambda^A + \lambda^p\left(a_2^B a_{31}^{*A} - a_1^B a_{32}^{*A}\right) - \lambda^L a_{33}^{*A}}{\lambda^A + \lambda^p\left(a_2^B a_{j1}^{*A} - a_1^B a_{j2}^{*A}\right) - \lambda^L a_{j3}^{*A}}, j = 1, 2. \tag{37}$$

Recall $\lambda^p = 0$. From the private equilibrium until the mimicking constraints bind λ^L is increasing from zero at the private equilibrium, so the RHS must be increasing and thus, with $t_j = 0$, t^H must be increasing from zero. Moving up the upf beyond the mimicking-constraints point, the Bs are getting a larger cash transfer and everyone is paying higher commodity taxes. If t^H stayed constant or rose there would be no reason for As not to mimic Bs. Therefore, after mimicking constraints bind, t^H must fall to prevent the As from mimicking the Bs.

Case 2, $a_{13}^{*j} > a_{23}^{*j}$: Using **R1**, from the private equilibrium until mimicking constraints bind, (34) tells us that as λ^L rises from zero λ^p falls below zero. Then using (29) and (30),

$t_1 < 0$ and $t_2 > 0$. And, if leisure is weakly separable from goods so that $x_j^{AB} = x_j^B, j = 1, 2$ the same argument says that $t_1 > t_2$ after mimicking constraints bind. If leisure is not weakly separable from goods and mimicking constraints bind could $t_1 = t_2$? Using **R1** and **R3** (33) would force λ^p to be positive. But then (32) would imply $t_1 > t_2$, which contradicts the assumption $t_1 = t_2$. What must happen is that with leisure not weakly separable from goods (34) implies λ^p is a bit less negative and then (33) implies $t_1 > t_2$. *Mutatis mutandis*, one can prove that if $a_{13}^{*j} < a_{23}^{*j}$, between the private equilibrium and the point where mimicking constraints bind, $t_1 < 0$ and $t_2 > 0$; and after mimicking constraints bind $t_1 < t_2$.

To complete the proof of Statement 2 I need to show that what was said of t^H in case 1 holds in case 2. The argument as to why t^H must fall along the upf after mimicking constraints bind is independent of the sign of $a_{13}^{*j} - a_{23}^{*j}$. All that remains is to explain why t^H must increase from zero at the private equilibrium until the mimicking constraints bind. Take the case where $a_{13}^{*j} > a_{23}^{*j}$. From the argument above we know $\lambda^p < 0, t_1 > 0$ and $t_2 < 0$. Look at (37) with $j = 2$. The λ^p terms in the numerator and denominator are both negative but using **R1** and **R3** the absolute value of the λ^p term in the numerator will be small relative to the λ^p term in the denominator. Therefore, the RHS of (37) must still increase. This means $(1 + t_2)/(1 - t^H)$ must increase and since t_2 is falling below zero t^H must increase.

References

[1] Atkinson, A. and J. Stiglitz, 1976, The design of tax structure: direct versus indirect taxation, *Journal of Public Economics* 6, 55–75.

[2] Auerbach, A.J., 1979, A brief note on a non-existent theorem about the optimality of uniform taxation, *Economics Letters* 3, 49–52.

[3] Burbidge, J.B., 2015, Using distance functions to understand interest taxation, *Canadian Journal of Economics* 48, pp. 903–923.

[4] Conesa, J.C., S. Kitao and D. Krueger, 2009, Taxing capital? Not a bad idea after all! *American Economic Review* 99, pp. 25–48.

[5] Corlett, W.J. and D.C. Hague, 1953-54, Complementarity and the excess burden of taxation, *The Review of Economic Studies* 21, pp. 21–30.

[6] Deaton, A.S., 1979, The distance function and consumer behaviour with applications to index numbers and optimal taxation, *Review of Economic Studies* 46, 391–405.

[7] Deaton, A.S., 1981, Optimal taxes and the structure of preferences, *Econometrica* 49, 391–405.

[8] Diamond, P.A., 1980, Income taxation with fixed hours of work, *Journal of Public Economics* 13, 101–110.

[9] Diamond, P.A. and J.A. Mirrlees, 1971, Optimal taxation and public production, *American Economic Review* 61, 8–27 and 261–278.

[10] Erosa, A. and M. Gervais, 2002, Optimal taxation in life-cycle economies, *Journal of Economic Theory* 105, 338–369.

[11] Jacobs, B. and R. Boadway, 2014, Optimal linear commodity taxation under optimal non-linear income taxation, *Journal of Public Economics* 117, 201–210.

[12] Mirrlees, J.A., 1971, An exploration in the theory of optimum income taxation, *The Review of Economic Studies* 38, 175–208.

[13] Piketty, T. and E. Saez, 2013, Optimal labor income taxation, in A. Auerbach, R. Chetty, M. Feldstein and E. Saez, editors, *Handbook of Public Economics*, vol. 5 (Amsterdam: Elsevier-North Holland), 391-474.

[14] Ramsey, F.P., 1927, A contribution to the theory of taxation, *Economic Journal* 37, 47–61.

[15] Saez, E., 2002, Optimal income transfer programs: intensive versus extensive labor supply responses, *Quarterly Journal of Economics* 117, 1039–1073.

Table 1**: An example of a Pareto efficient tax structure with four types - A,B,C,D; mimicking constraints bind in the order CD first, then AB, then BC; utility = x1^alpha*x2^beta*l^(1-alpha-beta)

Parameters

L=5	alpha = 3/10	beta = 9/20	wA=2	wB=1.7	wC=1.1	wD=1	nA=1	nB=1	nC=1	nD=1	p1=1	p2=1	tL=0

Line	ateH	ateU	ateM	ateL	tH	tU	tM	t1	t2
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.008081	0.000000	0.000000	-0.016228	0.008081	0.000000	0.000000	0.000000	0.000000
3	0.037926	0.000000	-0.027450	-0.046396	0.037926	0.000000	0.000000	0.000000	0.000000
4	0.042164	-0.000236	-0.031193	-0.050514	0.042164	0.042164	0.000000	0.000000	0.000000
5	0.112078	0.072772	-0.156740	-0.189142	0.112078	0.112078	0.000000	0.000000	0.000000
6	0.066682	0.025367	-0.231406	-0.272062	0.066682	0.066682	0.000000	0.056446	0.056446

Line	x1A	x2A	lA	utA	x1B	x2B	lB	utB	x1C	x2C	lC	utC	x1D	x2D	lD	utD
1	3.000000	4.500000	1.250000	2.892741	2.550000	3.825000	1.250000	2.560788	1.650000	2.475000	1.250000	1.847488	1.500000	2.250000	1.250000	1.720034
2	2.975756	4.463634	1.250000	2.875190	2.550000	3.825000	1.250000	2.560788	1.650000	2.475000	1.250000	1.847488	1.518183	2.277274	1.265152	1.740884
3	2.886221	4.329331	1.250000	2.810061	2.550000	3.825000	1.250000	2.560788	1.683738	2.525606	1.275559	1.885263	1.551597	2.327395	1.292997	1.779200
4	2.873508	4.310262	1.250000	2.800773	2.522685	3.784027	1.291046	2.560788	1.688303	2.532454	1.279017	1.890375	1.556119	2.334179	1.296766	1.784385
5	2.663766	3.995650	1.250000	2.646005	2.338550	3.507825	1.291046	2.419282	1.836652	2.754977	1.391403	2.056479	1.703178	2.554766	1.419315	1.953015
6	2.650352	3.975528	1.250000	2.636005	2.326774	3.490160	1.291046	2.410139	1.818081	2.727122	1.455079	2.063824	1.691122	2.536682	1.488815	1.965995

Line	x1AB	x2AB	lAB	utAB	x1BC	x2BC	lBC	utBC	x1CD	x2CD	lCD	utCD
1	2.550000	3.825000	1.812500	2.810061	1.650000	2.475000	2.573529	2.213025	1.500000	2.250000	1.590909	1.826926
2	2.550000	3.825000	1.812500	2.810061	1.650000	2.475000	2.573529	2.213025	1.518183	2.277274	1.604684	1.847488
3	2.550000	3.825000	1.812500	2.810061	1.683738	2.525606	2.590067	2.250477	1.551597	2.327395	1.629998	1.885263
4	2.522685	3.784027	1.847389	2.800773	1.688303	2.532454	2.592305	2.255539	1.556119	2.334179	1.633424	1.890375
5	2.338550	3.507825	1.847389	2.646005	1.836652	2.754977	2.665025	2.419282	1.703178	2.554766	1.744831	2.056479
6	2.326774	3.490160	1.847389	2.636005	1.818081	2.727122	2.706227	2.410139	1.691122	2.536682	1.808014	2.063824

Line	lamA	lamB	lamC	lamD	lampA	lampB	lampC	lamL	lamAB	lamBC	lamCD
1	10.000000	8.500000	5.500000	5.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	9.939390	8.500000	5.500000	5.060610	0.000000	0.000000	0.000000	0.101016	0.000000	0.000000	0.000000
3	9.715551	8.500000	5.612459	5.171990	0.000000	0.000000	0.000000	0.474081	0.000000	0.000000	0.000000
4	9.683770	9.320728	5.627676	5.187064	0.000000	0.000000	0.000000	0.527050	0.911779	0.000000	0.000000
5	9.159416	10.218804	6.122172	5.677258	0.000000	0.000000	0.000000	1.400973	2.423637	0.000000	0.000000
6	9.125880	10.276242	6.584854	6.902708	0.000000	0.000000	0.000000	1.456867	2.520330	0.524584	1.265636

Line 1: Private equilibrium; all taxes zero
Line 2: utA is reduced until CD mimicking constraints start to bind; note tM = 0; after this utC is tied to utD
Line 3: utA further reduced until AB mimicking constraints start to bind
Line 4: utA further reduced with tU rising until it equals tH to cope with AB mimicking constraints
Line 5: utA reduced until BC mimicking constraints start to bind
Line 6: utA reduced by 0.01; CD, AB and BC mimicking constraints bind

Notation
ate: average earnings tax rate
t1, t2: commodity tax rates
tj, j = H,U,M,L; marginal earnings tax rates on high, upper middle, middle and low wage earners
lamj, j = A,B,C,D; Lagrange multipliers on utility levels for A, B, C and D
lampj, j = A,B,C; Lagrange multipliers on same commodity prices for A and D, B and D, and C and D
lamL; Lagrange multiplier on 1/L-a3A ge 0
lamAB; Langrange multiplier on utA ge utAB
lamBC; Langrange multiplier on utB ge utBC
lamCD; Langrange multiplier on utC ge utCD

**** The programs used to produce these simulations are written in R; I thank Pierre Chaussé of the University of Waterloo for assistance.**

Table 2: An example of a Pareto efficient tax structure with four types - A,B,C,D; mimicking constraints bind in the order CD first, then AB, then BC; utility = (x1*x2)^(1/2) + x1^(1/2) + l^(1/2)

Parameters

L=5	wA=2	wB=1.9	wC=1.1	wD=1	nA=1	nB=1	nC=1	nD=1	p1=1	p2=1	tL=0

Line	ateH	ateU	ateM	ateL	tH	tU	tM	t1	t2
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.010051	-0.000001	-0.000002	-0.021882	0.010051	0.000001	1.463 E-07	0.000017	-0.000023
3	0.010051	-0.000001	-0.000002	-0.021882	0.010051	0.000001	0.000000	0.000017	-0.000023
4	0.012353	-0.000001	-0.002238	-0.024385	0.012353	0.000001	0.000000	0.000021	-0.000029
5	0.012696	-0.000008	-0.002566	-0.024752	0.012696	0.012696	0.000000	0.000022	-0.000030
6	0.141937	0.132322	-0.248507	-0.300731	0.141937	0.141937	0.000000	0.000337	-0.000447
7	0.068567	0.058191	-0.376097	-0.444643	0.068567	0.068567	0.000000	0.092288	0.091383

Line	x1A	x2A	lA	utA	x1B	x2B	lB	utB	x1C	x2C	lC	utC	x1D	x2D	lD	utD
1	5.815822	3.852933	0.165623	7.552275	5.529488	3.625431	0.181622	7.255020	3.162780	1.815409	0.474373	4.863361	2.851708	1.590541	0.557750	4.565259
2	5.758271	3.807278	0.168662	7.492561	5.529389	3.625530	0.181622	7.255020	3.162727	1.815463	0.474373	4.863361	2.906464	1.629904	0.560760	4.630199
3	5.758271	3.807278	0.168662	7.492561	5.529389	3.625530	0.181622	7.255020	3.162727	1.815463	0.474373	4.863361	2.906464	1.629904	0.560760	4.630199
4	5.745081	3.796823	0.169370	7.478880	5.529366	3.625553	0.181622	7.255020	3.168986	1.820051	0.474627	4.870703	2.912714	1.634416	0.561099	4.637610
5	5.743115	3.795265	0.169476	7.476841	5.524447	3.621663	0.186288	7.255020	3.169902	1.820723	0.474665	4.871779	2.913629	1.635077	0.561149	4.638696
6	4.997017	3.209445	0.217887	6.706889	4.795815	3.052451	0.239250	6.505160	3.851988	2.328928	0.499226	5.664371	3.594961	2.136205	0.593691	5.437756
7	4.967758	3.186734	0.220179	6.676889	4.767275	3.030418	0.241755	6.475994	3.810516	2.297820	0.593610	5.681549	3.565837	2.114620	0.706358	5.474771

Line	x1AB	x2AB	lAB	utAB	x1BC	x2BC	lBC	utBC	x1CD	x2CD	lCD	utCD
1	5.529488	3.625431	0.422541	7.478880	3.162780	1.815409	2.379900	5.717306	2.851708	1.590541	0.961591	4.799040
2	5.529389	3.625530	0.422541	7.478880	3.162727	1.815463	2.379900	5.717306	2.906464	1.629904	0.964327	4.863361
3	5.529389	3.625530	0.422541	7.478880	3.162727	1.815463	2.379900	5.717306	2.906464	1.629904	0.964327	4.863361
4	5.529366	3.625553	0.422541	7.478880	3.168986	1.820051	2.380047	5.724512	2.912714	1.634416	0.964636	4.870703
5	5.524447	3.621663	0.426973	7.476841	3.169902	1.820723	2.380069	5.725567	2.913629	1.635077	0.964681	4.871779
6	4.795815	3.052451	0.477287	6.706889	3.851988	2.328928	2.394289	6.505160	3.594961	2.136205	0.994264	5.664371
7	4.767275	3.030418	0.479667	6.676889	3.810516	2.297820	2.448932	6.475994	3.565837	2.114620	1.096690	5.681549

Line	lamA	lamB	lamC	lamD	lampA	lampB	lampC	lamL	lamAB	lamBC	lamCD
1	10.000000	9.500000	5.500000	5.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	9.902872	9.500000	5.500000	5.097128	-0.010167	0.004776	0.001509	0.034855	0.000000	0.000000	0.000000
3	9.902872	9.500000	5.500000	5.097130	-0.010167	0.004776	0.001509	0.034855	0.000000	0.000000	0.000003
4	9.880643	9.500000	5.511127	5.108233	-0.012573	0.005912	0.001876	0.043023	0.000000	0.000000	0.000003
5	9.877331	9.823343	5.512756	5.109859	-0.012932	0.006077	0.001932	0.044246	0.333979	0.000000	0.000003
6	8.642237	12.400667	6.730065	6.324900	-0.225138	0.085759	0.055269	0.640808	4.243618	0.000000	0.000045
7	8.594850	12.523159	7.568399	8.477857	-0.234622	0.089578	0.057561	0.670705	4.418347	0.941308	2.204585

Line 1: Private equilibrium; all taxes zero
Line 2: utA is reduced until CD mimicking constraints start to bind; note tM is a small positive number
Line 3: In this line utA is reduced by enough with utC constant to make tM zero; after this, utC is tied to utD
Line 4: utA further reduced until AB mimicking constraints start to bind
Line 5: utA further reduced with tU rising until it equals tH to cope with AB mimicking constraints
Line 6: utA reduced until BC mimicking constraints start to bind
Line 7: utA reduced by 0.03; CD, AB and BC mimicking constraints bind

Notation
ate: average earnings tax rate
t1, t2: commodity tax rates
tj, j = H,U,M,L; marginal earnings tax rates on high, upper middle, middle and low wage earners
lamj, j = A,B,C,D; Lagrange multipliers on utility levels for A, B, C and D
lampj, j = A,B,C; Lagrange multipliers on same commodity prices for A and D, B and D, and C and D
lamL; Lagrange multiplier on 1/L-a3A ge 0
lamAB; Langrange multiplier on utA ge utAB
lamBC; Langrange multiplier on utB ge utBC
lamCD; Langrange multiplier on utC ge utCD

Table 3: An example of a Pareto efficient tax structure with four types - A,B,C,D; mimicking constraints bind in the order CD first, then AB, then BC; utility = (x1*x2)^(1/2) + (x1*l)^(1/2)

Parameters

L=5	wA=2	wB=1.9	wC=1.1	wD=1	nA=1	nB=1	nC=1	nD=1	p1=1	p2=1	tL=0

Line	ateH	ateU	ateM	ateL	tH	tU	tM	t1	t2
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.006128	-0.000135	-0.000165	-0.014115	0.006128	0.000038	0.000007	0.000387	-0.000258
3	0.006121	-0.000142	-0.000172	-0.014123	0.006121	0.000031	0.000000	0.000394	-0.000251
4	0.017352	-0.000405	-0.011931	-0.027285	0.017352	0.000087	0.000000	0.001125	-0.000715
5	0.018549	-0.000834	-0.013862	-0.029450	0.018549	0.018549	0.000000	0.002146	-0.001360
6	0.101461	0.084145	-0.187917	-0.225920	0.101461	0.101461	0.000000	0.012383	-0.007696
7	0.056400	0.038228	-0.271564	-0.321403	0.056400	0.056400	0.000000	0.073135	0.044381

Line	x1A	x2A	lA	utA	x1B	x2B	lB	utB	x1C	x2C	lC	utC	x1D	x2D	lD	utD
1	5.000000	3.333333	0.833333	6.123724	4.500000	2.892857	0.892857	5.612486	2.750000	1.440476	1.190476	3.799671	2.500000	1.250000	1.250000	3.535534
2	4.967439	3.307249	0.836608	6.091794	4.498727	2.894130	0.892857	5.612486	2.749278	1.441198	1.190476	3.799671	2.525396	1.263675	1.263024	3.572371
3	4.967437	3.307248	0.836608	6.091792	4.498726	2.894131	0.892857	5.612486	2.749278	1.441199	1.190476	3.799671	2.525393	1.263669	1.263036	3.572372
4	4.907716	3.259439	0.842687	6.033184	4.496367	2.896492	0.892858	5.612486	2.771812	1.455070	1.200819	3.832681	2.547953	1.276778	1.274954	3.606020
5	4.896748	3.256944	0.843007	6.025295	4.478139	2.870951	0.917406	5.612486	2.773006	1.458570	1.202154	3.836935	2.549358	1.280024	1.276545	3.610431
6	4.437739	2.916897	0.889358	5.584477	4.056073	2.564656	0.965384	5.204084	3.089156	1.656945	1.348380	4.303344	2.866061	1.467676	1.445172	4.086141
7	4.396463	2.908138	0.890629	5.554477	4.018286	2.556778	0.966696	5.176191	3.056372	1.610989	1.452196	4.325722	2.844824	1.429846	1.559578	4.123197

Line	x1AB	x2AB	lAB	utAB	x1BC	x2BC	lBC	utBC	x1CD	x2CD	lCD	utCD
1	4.643073	2.749784	1.303571	6.033355	2.990532	1.199945	2.671958	4.721087	2.562300	1.187700	1.590909	3.763496
2	4.641744	2.751021	1.303571	6.033295	2.989731	1.200590	2.671958	4.720964	2.587530	1.201501	1.602749	3.799671
3	4.641743	2.751022	1.303571	6.033295	2.989731	1.200590	2.671958	4.720964	2.587526	1.201496	1.602760	3.799671
4	4.639280	2.753317	1.303572	6.033184	3.012143	1.214296	2.678278	4.752804	2.609913	1.214704	1.613595	3.832681
5	4.617568	2.731033	1.325666	6.025295	3.013137	1.217597	2.679094	4.756618	2.611246	1.217919	1.615041	3.836935
6	4.180220	2.437996	1.368845	5.584477	3.326529	1.414770	2.768454	5.204084	2.925389	1.407147	1.768338	4.303344
7	4.141217	2.430462	1.370027	5.554477	3.275925	1.385391	2.831897	5.176191	2.898773	1.374411	1.872344	4.325722

Line	lamA	lamB	lamC	lamD	lampA	lampB	lampC	lamL	lamAB	lamBC	lamCD
1	10.000000	9.000000	5.500000	5.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	9.947904	9.000000	5.500000	5.052095	-0.132867	0.057890	0.020064	0.091595	0.000000	0.000000	0.000000
3	9.947902	9.000000	5.500000	5.052258	-0.132835	0.057922	0.020075	0.091598	0.000000	0.000000	0.000174
4	9.852529	9.000003	5.547783	5.100151	-0.380066	0.166864	0.058848	0.261327	-0.000001	0.000000	0.000504
5	9.839706	9.593500	5.553946	5.106816	-0.236335	0.315574	0.112560	0.278210	0.642861	0.000000	0.000962
6	9.133353	11.684322	6.229318	5.785152	-1.591946	1.687959	0.921472	1.596074	3.608175	-0.000002	0.006699
7	9.085860	11.781650	6.761463	7.482956	-1.023588	2.252718	1.215187	1.663789	3.760825	0.622428	1.764030

Line 1: Private equilibrium; all taxes zero
Line 2: utA is reduced until CD mimicking constraints start to bind; note tM is a small positive number
Line 3: In this line utA is reduced by enough with utC constant to make tM zero; after this, utC is tied to utD
Line 4: utA further reduced until AB mimicking constraints start to bind
Line 5: utA further reduced with tU rising until it equals tH to cope with AB mimicking constraints
Line 6: utA reduced until BC mimicking constraints start to bind
Line 7: utA reduced by 0.03; CD, AB and BC mimicking constraints bind

Notation
ate: average earnings tax rate
t1, t2: commodity tax rates
tj, j = H,U,M,L; marginal earnings tax rates on high, upper middle, middle and low wage earners
lamj, j = A,B,C,D; Lagrange multipliers on utility levels for A, B, C and D
lampj, j = A,B,C; Lagrange multipliers on same commodity prices for A and D, B and D, and C and D
lamL; Lagrange multiplier on 1/L-a3A ge 0
lamAB; Langrange multiplier on utA ge utAB
lamBC; Langrange multiplier on utB ge utBC
lamCD; Langrange multiplier on utC ge utCD