

REVISITING THE CLASSICAL VIEW OF BENEFIT-BASED TAXATION*

Matthew Weinzierl

This article incorporates into modern optimal tax theory the classical logic of benefit-based taxation in which an individual's benefit from the activities of the state is tied to his or her income-earning ability. First-best optimal policy is characterised analytically as depending on a few potentially estimable statistics. Constrained optimal policy, with a Pareto-efficient objective that trades off this principle and conventional utilitarianism, is simulated using conventional constraints and methods. A wide range of optimal policy outcomes can result, including those that match well several features of existing policies. This analysis thereby contributes to the theory of positive optimal taxation.

In February, 1913, the Sixteenth Amendment to the US Constitution was passed, allowing for the direct taxation of incomes. One month later, the nascent *American Economic Review* included a paper by Blakey (1914), a noted tax expert and University of Minnesota professor, that gave the following argument for the propriety of taxing large (mainly urban) incomes:

These [urban] centers and their wealthy residents have cause for satisfaction and thanksgiving that their incomes are so bountiful and that the country has provided them with such great opportunities, rather than occasion for criticising the requirement of a moderate contribution to the nation which has rendered such incomes possible.

In 1935, US President Franklin Delano Roosevelt sought to increase the progressivity of taxes substantially (Roosevelt, 1935). He argued as follows:

With the enactment of the Income Tax Law of 1913, the Federal Government began to apply effectively the widely accepted principle that taxes should be levied in proportion to ability to pay and in proportion to the benefits received. Income was wisely chosen as the measure of benefits and of ability to pay. This was, and still is, a wholesome guide for national policy. It should be retained as the governing principle of Federal taxation.

In 2011, US President Barack Obama also sought to increase top marginal income tax rates (Obama, 2011). He applied this reasoning:

As a country that values fairness, wealthier individuals have traditionally borne a greater share of this [tax] burden than the middle class or those less fortunate. Everybody pays, but the wealthier have borne a little more. This is

* Corresponding author: Matthew Weinzierl, Harvard Business School, Morgan Hall 277, Boston, MA 02163, USA. Email: mweinzierl@hbs.edu.

Thanks to the editor, Frederic Vermeulen, the referees and Alan Auerbach, Rafael Di Tella, Roger Gordon, Louis Kaplow, Wojciech Kopczuk, Ilyana Kuziemko, Benjamin B. Lockwood, Greg Mankiw, David Moss, Kristi Olson, Andreas Peichl, Florian Scheuer, Daniel Shaviro, Pierre Yared, Glen Weyl and participants at the 2013 NTA conference, NYU Law School Tax Policy Colloquium, Columbia University and the University of Chicago's Becker Friedman Institute's inaugural Normative Ethics and Welfare Economics conference.

not because we begrudge those who've done well – we rightly celebrate their success. Instead, it is a basic reflection of our belief that those who've benefited most from our way of life can afford to give back a little bit more.

Modern tax theorists will find the normative arguments underlying these quotations both familiar and strange. All three refer to the logic of 'benefit-based taxation', under which people ought to pay taxes that depend on how much they benefit from public goods. All three quotes also refer to the logic of 'ability-based taxation', under which people ought to pay taxes that depend on how much they are hurt by having to earn the money to pay. Introductory public finance textbooks describe these two ideas as the classic principles of optimal tax design. But they are usually considered alternatives, not complements. To modern tax theorists working in the tradition of Mirrlees (1971), stranger still is the use of these principles at all, as modern tax theory has largely moved away from them in favour of an approach emphasising social welfare maximisation.

Two hundred years ago, however, the normative principle underlying these statements would have been recognised immediately as Smith's (1776 [1991]) first maxim of taxation:

The subjects of every state ought to contribute toward the support of the government, as near as possible, in proportion to their respective abilities; that is in proportion to the revenue which they respectively enjoy under the protection of the state. (1991, p. 498)

Like those quoted above, Smith argues that one's income – as a measure of one's ability to pay – is a measure of one's benefit from the state. Because he supports benefit-based taxation more generally, Smith believes that the ability to pay is therefore an appropriate basis for taxation.¹ Simply put, Smith endorses benefit-as-ability-based taxation. Musgrave (1959) labelled this logic the 'classical' view of benefit-based taxation, a label I adopt in this article.

This classical view of benefit-based taxation was highly influential in the late eighteenth and early nineteenth centuries but waned as Mill's (1871 [1994]) purely ability-based reasoning gained favour and the canonical contributions of Lindahl (1919) shifted the focus of benefit-based tax research.² The idea of benefit-as-ability was not further explored, while benefit-based and ability-based reasoning were developed as separate ideas. Benefit-based reasoning was assigned a subsidiary role in tax theory, namely as a means by which to value and assign the costs of public expenditures while, crucially, taking the distribution of income as given.³ (Note that the classical view has endogeneity of the income distribution as its core component.) Ability-based reasoning was absorbed into the now-dominant Mirrleesian approach, as Mirrlees (1971) made differences in the ability to earn income the linchpin of taxation

¹ Smith argued that the preferred way of paying for public goods was to have them privately or locally managed, with user fees mechanically tying the funding for these institutions to individual benefits, though he recognised user fees were infeasible in many cases. See Smith (1776 [1991], Book V Article 1)

² As the remarkable surveys by Seligman (1908) and Musgrave (1959) make clear, benefit-based reasoning was a prominent, at times leading, approach among tax theorists through the nineteenth century. Petty (1677 cited in Musgrave, 1959.) anticipated Smith's view, and Hobbes, Hume, and Rousseau among others subscribed to it in some form.

³ Lindahl himself viewed his theory as conditional in this way, a view criticised as untenable by Samuelson (1955).

in his theory.⁴ The classical benefit-based logic exerted little influence on the welfarist objective assumed in modern Mirrleesian theory.

In normative terms, the shift from the classical benefit-based view to the dominant modern approach, which pursues the so-called ‘endowment taxation’, is quite substantial. Under the modern approach, an individual’s income-earning ability is taken as a given and, as ability makes it easier to obtain consumption in exchange for leisure, social welfare maximisation will tend to imply taxing those with higher endowments to supplement the resources of others. In contrast, the classical benefit-based view treats an individual’s ability as a function of the activities of the state to which that individual contributes. Any endowments are moot unless individuals cooperate to fund public goods that set the stage for abilities to develop. Under this approach, taxation is merely the means by which individuals pay for these public goods cooperatively, each according to how much he or she values them.

Perhaps due in part to this normative contrast, the modern shift away from using the benefit-based approach, in any form, as a general principle of taxation appears nearly complete. The authoritative review of modern tax theory entitled the Mirrlees Review (Institute of Fiscal Studies *et al.*, 2010) makes no reference to Lindahl’s (1919) canonical development of a benefit-based theory or to any of the relatively few more recent refinements of it. Atkinson and Stiglitz’s (1980) classic text and Kaplow’s (2008) invaluable modern treatise each devote only a few pages among hundreds to benefit-based taxation, the latter largely to point out its weaknesses (as discussed below). Neither Salani’e (2011), in his essential textbook on the economics of taxation, nor Boadway (2012), in his excellent survey of optimal tax theory’s implications for policy, mention Lindahl or benefit-based taxes.⁵

The long-standing role for classical benefit-based logic in public reasoning over taxes stands in stark contrast to this momentum away from it in modern theory. The purpose of this article is to explore whether we might reconcile this disconnect by incorporating the classical view of benefit-based taxation into the modern framework of optimal tax theory, thereby resuscitating it as part of how we understand policy design.

The first contribution of this article is the finding that the classical benefit-based view can fit neatly into the Mirrleesian approach once one makes a simple – and arguably needed – change to the standard setup: that is, allowing individual income-earning ability to be a function of both endowed ability and public goods. Once public goods matter for ability, the classical benefit-as-ability view links seamlessly with the modern model. Thus, both first-best and constrained optimal benefit-based policy can be analysed within the formal structure of modern tax theory and characterised using familiar methods.⁶ Of course, the normative contrast between the classical benefit-based view and endowment taxation remains.

⁴ The Mirrleesian approach’s linkage of ability to pay and ability to earn relies on its assumption – Mirrlees (1971) makes it his second assumption – that tastes are homogenous. See Lockwood and Weinzierl (2015).

⁵ As noted, benefit-based reasoning continues to occupy a prominent but narrow role in studies of public goods provision that take the income distribution as given. See, for example, Aaron and McGuire (1970) and the large literature following upon their work, as cited below.

⁶ One way to interpret this contribution is that it shows how this classical view might – by linking benefit to ability – avoid the common critique of benefit-based reasoning that it ‘has little more than emotive content’ and ‘leads nowhere at all’, as Simons (1938) put it (quoted in Shaviro, 2013).

In particular, first-best policy according to this view of benefit-based taxation can be characterised in terms of simple and potentially observable elasticity parameters if we apply Lindahl's well-known method of measuring benefit. In this article, we derive the version of the Samuelson rule for the optimal extent of public goods under the classical benefit-based view. We also obtain a straightforward condition determining the progressivity of optimal average tax rates, which turns out to depend in an intuitive way on the size of the Hicksian coefficient of complementarity between public goods and endowed ability. These conditions reduce to especially simple relationships if we assume that the ability production function takes certain forms. As an interesting aside, assuming those same forms and a familiar form for the individual utility function, we find that optimal benefit-as-ability-based taxation quantitatively resembles Mill's preferred 'equal sacrifice' taxation, a possibility hinted at informally over forty years ago by Feldstein (1976).

As with the first-best policy, we find that the classical view's linkage of benefit and ability facilitates the analysis of constrained optimal benefit-based policy. We need to modify only the objective, not the constraints, of conventional Mirrleesian analysis to characterise how different assumptions about the interaction of public goods and endowed ability affect the progressivity of constrained optimal benefit-based policy. To address the limitation that the classical benefit-based view does not provide a ranking of allocations other than the first-best, we modify the standard conventional utilitarian objective function to allow the planner to choose the allocation, from the set of incentive compatible allocations, that deviates least from the first-best benefit-based allocation while respecting Pareto efficiency.

Why go through the trouble of attempting such a resuscitation? The statements provided at the start of this article suggest that the classical view of benefit-based taxation is included in the criteria used to judge, or at least justify, tax policy in the US. If one takes a positive approach to specifying the objective of optimal taxation, that view therefore ought to be included in our models, as well.⁷ While conventional optimal tax analysis uses an objective for policy based on philosophical reasoning, recent work has pursued the idea of basing that objective on evidence of the normative priorities that prevail in society.⁸ An important feature of this 'positive optimal tax theory' is its inclusion of multiple normative criteria, as a wide range of evidence has shown that most persons base their moral judgments on more than one principle.⁹

⁷ Some readers may question the usefulness of popular conceptions of economic justice for normative policy analysis. The perspective implicit in this article is that a back-and-forth between introspective reasoning on moral principles and robust popular reasoning is likely to be productive. See Sheffrin (2013) for the development of the related idea of a role for 'folk justice' and Gaertner and Schokkaert (2012) for related arguments behind the use of 'empirical social choice' methods.

⁸ See Saez and Stantcheva (2015) and Weinzierl (2014); also see Zelenak (2006) for a different, legal theory, perspective and Lockwood and Weinzierl (2016) for related evidence. Such a positive optimal tax theory differs from a purely positive theory of taxation, where the political process and self-interest of voters play central roles. Positive optimal tax theory retains the conventional theory's focus on an objective that reflects the moral reasoning of an impartial observer.

⁹ See the discussion of normative diversity in Weinzierl (2014). Not mentioned there, however, is that Adam Smith appears to have had a similarly mixed perspective. While in his first maxim, he seems to argue for proportional taxation, he writes at another point in the same book: 'The necessities of life occasion the great expence of the poor. They find it difficult to get food, and the greater part of their little revenue is spent in getting it. The luxuries and vanities of life occasion the principal expence of the rich . . . It is not very unreasonable that the rich should contribute to the public expence, not only in proportion to their revenue, but something more than in that proportion'. (p. 513)

The second contribution of this article is, therefore, to explore whether an optimal tax model with a mixed objective function that gives weight to both the classical view of benefit-based taxation and the conventional utilitarian criterion has quantitative explanatory power. To implement that analysis, I take advantage of a convenient feature of the objective function used in the preceding first-best analysis. That objective function differs from the conventional utilitarian criterion for optimal policy in only one way: it treats gains and losses of utility from a benchmark allocation asymmetrically, such that the utilitarian criterion is equivalent to restoring symmetry to that objective function. Moreover, that asymmetry is controlled by a single parameter that I label δ and which provides a simple way to adjust the relative weights given to the classical benefit-based criterion and the conventional utilitarian criterion. By varying δ , we find that a wide range of optimal policy outcomes can result, including those consistent with existing policy. These results suggest that this model may offer a useful approach to a positive optimal tax theory.

Readers familiar with Weinzierl (2014) will find these two papers' analyses complementary. In that paper, I demonstrate public affinity for a different (but as noted below, related) principle of tax design: equal sacrifice. I then show that a role for equal sacrifice in the optimal tax objective function yields policy implications that resolve a number of puzzling gaps between conventional theory and robust features of prevailing tax policy. I leverage a similar, though streamlined, technique in Section 3 of this article to explore the policy implications of classical benefit-based taxation, which not surprisingly resemble those of equal sacrifice in important ways. Together, these two papers formalise and characterise the implications of including in the policy objective two long-standing and popular normative approaches to taxation that depart from the conventional approach in optimal tax theory. In so doing, these papers can help in understanding why actual tax policy looks the way it does.

Before proceeding, it may be important to highlight a few potentially related things that this article is not trying to do.

First, this article is not intended to defend classical benefit-based taxation as a normative principle. Other principles, including utilitarianism, may well be preferable from the perspective of moral and political philosophy. Instead, this article is intended to capture the view of benefit-based taxation that arose in early sophisticated thinking about optimal taxation and that appears to have retained a prominent role in public reasoning over taxes despite the strong criticisms of more recent tax theorists.

Second, this article does not claim to resolve the debate over whether the right approach to linking taxes to benefit is Lindahl's. The proper way to turn the normative intuition behind benefit-based taxation into individual tax burdens has long been a topic of study (prominent examples include Aaron and McGuire, 1970, 1976; Brennan, 1976*a,b*; West and Staaf, 1979; Moulin, 1987; Hines, 2000).¹⁰ Instead, I view Lindahl's

¹⁰ One of many disagreements from that debate is over the justifiability of using of an individual's marginal willingness to pay for public goods to measure that individual's benefit. The Lindahl answer to that question, as stated by Aaron and McGuire (1970), criticised by Brennan (1976*a,b*) and discussed by West and Staaf (1979), is to have two sets of taxes, one of which may offset differences in inframarginal benefits (as well as other undesirable inequalities) and the second of which implements the Lindahl benefit-based equilibrium. I adopt a version of that approach by having the objective of the social planner give weight to both the distribution of welfare (in the conventional utilitarian sense) and to the allocation's proximity to the first-best benefit-based allocation. Note that this article thereby aligns with the arguments of Samuelson (1954, footnote 9). An alternative approach would be to tie taxes to the total, inframarginal benefits from public goods, perhaps along the lines of Moulin (1987).

approach as a natural way to formalise the informal notion – implicit in rhetorical appeals to benefit-based taxation – that people ought to ‘pay for what they get’ from public goods just as they do in the market for private goods. The best known example of this metaphor is usually attributed to US Supreme Court Justice Oliver Wendell Holmes, who served on the Court from 1902 to 1932: ‘I like to pay taxes. With them I buy civilization’.¹¹ In fact, Lindahl (1919) entitled the full statement of his approach ‘Just Taxation – A Positive Solution’, and described it as a model of how ‘the distribution of the total cost of the collective goods ... is to be solved by free agreement ... as a kind of economic exchange’. (p. 168)

Third, the mixed normative objective utilised in this article is not meant as an alternative to the most general forms of modern optimal tax theory. As has been made clear in work by Stiglitz (1987) and Werning (2007), among others, Mirrleesian optimal tax theory imposes only the requirement of Pareto efficiency on the set of feasible and incentive compatible allocations. The objectives I consider satisfy Pareto efficiency as well, so in principle they could be modelled in alternative ways that are closer – in formal terms – to the standard approach (i.e. following the technique of Saez and Stantcheva, 2015). This article’s approach chooses among Pareto-efficient allocations by assigning weights to a small number of normative principles with apparent empirical relevance.

The article proceeds as follows. Section 1 formalises the classical benefit-based view and analytically characterises optimal policy under it by applying Lindahl’s methodology to an (otherwise standard) optimal tax model in which ability depends on public goods as well as endowed ability. Section 2 quantitatively characterises constrained-optimal (second-best) policy, again using the conventional modern optimal tax model modified only in its objective, which in this case is to deviate as little as possible from the allocation preferred by the classical benefit-based view. Section 3 extends that constrained-optimal analysis to the case of an objective function that can be interpreted as giving weight to both utilitarianism and the classical benefit-based view. Section 4 concludes. The Appendix contains additional examples of public rhetoric drawing on the classical benefit-based view as well as the proofs of analytical results.

1. Analytical Characterisation of Classical Benefit-based Taxation

In this Section, I show how a single modification of the standard Mirrleesian optimal tax model enables us to formalise classical benefit-based taxation (hereafter, CBBT) and characterise optimal policy in terms of simple, potentially observable statistics.

1.1. *The Model*

A population of individuals differ in their abilities to earn income. In conventional models, these abilities are interpreted as innate and fixed relative to public goods.¹²

¹¹ Attributed to Holmes in Frankfurter and MacLeish (1939). Holmes wrote a similar phrase into his decision on a 1927 case before the Court. It is worth recalling that, at the time, the US income tax was highly concentrated on only the wealthiest Americans, so Holmes was effectively arguing on behalf of progressive taxation. The wide use of his dictum in popular writings on taxation over the intervening decades suggests that it has broad appeal.

¹² Much recent work in the Mirrleesian literature has considered other sources of endogeneity, such as human capital investment. See Kapicka (2006), Best and Kleven (2013), Stantcheva (2015), Gelber and Weinzierl (2016), among others.

The innovation in this article is that individuals' income-earning abilities are functions of both endowed ability and public goods funded by tax revenue.¹³ As mutually exclusive categories, endowed ability and public goods can be thought of as, respectively, the genetic or other endowments not affected by the size and scope of the state and the broad set of activities with which the state supports the functioning of an economy. Note that this article takes only the first, most basic, step towards modelling the effect of these public goods on individuals' abilities by leaving unexplored their composition. Future research may show how the mix, design and implementation of public goods matter, as well.¹⁴

Formally, income-earning ability $w_i = f(a_i, G)$, where $i \in I$ indexes endowed ability types a_i , $G \geq 0$ is the level of spending on public goods and $f(\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a differentiable ability production function. The conventional model is a special case in which $f(a_i, G) = a_i$ for all $i \in I$ and any G .

An individual of type i derives utility according to:

$$U(c_i, l_i) = u \left[c_i, \frac{y_i}{f(a_i, G)} \right], \quad (1)$$

where c_i is private consumption for individual i and y_i is i 's income, so that y_i/w_i is work effort. Denote partial derivatives of this utility function with respect to consumption and effort with $u_c[c_i, y_i/f(a_i, G)]$ and $u_l[c_i, y_i/f(a_i, G)]$. Note that this utility function does not assume additive separability between the utility from consumption and the disutility from effort, a simplifying assumption often made in optimal tax models.

A social planner chooses a tax system, including an optimal G^* . Individuals take that system as given and maximise their own utility, yielding equilibrium consumption and income allocations $\{c_i^*, y_i^*\}_{i=1}^I$ and utility levels:

$$U_i^* = u \left[c_i^*, \frac{y_i^*}{f(a_i, G^*)} \right], \quad (2)$$

for all $i \in I$. Note that there may be multiple optimal policy equilibria, as discussed below.

At this point, it may be worth noting that the simple change of making income-earning ability depend on public goods raises a number of questions for conventional Mirrleesian optimal taxation that could be pursued in future research.

¹³ An innovative paper by Matsumoto (2001) is the only other analysis of which I am aware in which public goods augment individual abilities in an optimal tax model. His focus is on how public goods can thereby relax incentive constraints.

¹⁴ Much of public spending in advanced economies is for social insurance rather than 'productive' public goods, a distinction this article's model can accommodate in a number of ways. First, in the static setting of this article, the value of social insurance to welfare may be captured by the policy maker's weight on the utilitarian objective discussed in Section 3, while the ability production function $f(a, G)$ may be specified to reflect the effect of public goods on ability including all forms of government spending (i.e. G may have weaker marginal effects if it is made up largely of social insurance). Alternatively, in a dynamic model with shocks that may be insured against (perhaps through government activities), social insurance spending may fit into a benefit-based framework more directly, though not through its effects on ability, so that the model would distinguish between the types of G and how they generate benefit.

1.2. *First-best Classical Benefit-based Taxes*

With the endogenous ability defined above, we can now apply Lindahl's theory to determine optimal benefit-based taxes. The core insight of Lindahl's work starts with the idea of personalised tax 'prices', or shares.¹⁵ The planner specifies this set of shares $\{\tau_i\}$ and individual i therefore faces the budget constraint:

$$y_i - c_i - \tau_i p G \geq 0, \quad (3)$$

where p is the per-unit cost of public goods relative to private consumption goods and τ_i is the share of the total cost of public goods paid for by the individuals of type i .¹⁶

Lindahl's approach as stated here is a 'first-best' one, as it assumes that individual type is observable; i.e. the tax planner can assign personalised shares τ_i to individuals. This unrealistic assumption has long been a target for criticism of benefit-based taxation in the Lindahl tradition. In the next Section, I show how the classical logic for benefit-based taxes – due to its treatment of ability as benefit – allows us to accommodate incentive compatibility within the Lindahl approach just as the Mirrlees model does.

In this subsection, I follow Lindahl and focus on first-best benefit-based policies. This is not merely for analytical tractability, though that is a side benefit. Rather, first-best policies are the most direct guide to a normative principle's effects on policy, as evidenced by the analytical results below and how they compare to first-best policies under a conventional utilitarian criterion. More concretely, first-best policies are essential ingredients in the calculation of constrained (second-best) optimal policies in the positive optimal tax model that I analyse numerically in the next Section.

Lindahl's approach then has us consider a hypothetical scenario in which each individual i is allowed to choose his own consumption, work effort, and, importantly, level of public goods provision that maximise his utility subject to his personal budget constraint, taking the tax share τ_i as given. Denote this individually-optimal level G_i and let λ denote the multiplier on the budget constraint (3). The individual first order conditions are as follows:

$$u_c \left[c_i, \frac{y_i}{f(a_i, G)} \right] = \lambda, \quad (4)$$

$$\frac{1}{f(a_i, G_i)} u_l \left[c_i, \frac{y_i}{f(a_i, G)} \right] = -\lambda, \quad (5)$$

$$\frac{-y_i f_G(a_i, G_i)}{f(a_i, G_i)} \frac{1}{f(a_i, G_i)} u_l \left[c_i, \frac{y_i}{f(a_i, G)} \right] = \tau_i p \lambda. \quad (6)$$

Lindahl defined optimal policy as that in which two conditions are satisfied: first, the personalised shares cause each type to prefer the same quantity of public goods;

¹⁵ I assume the public goods are entirely non-excludable. Taxing the benefits from excludable public goods raises additional possibilities: see for example Hellwig (2005).

¹⁶ I am restricting attention to what Kaneko (1977*a,b*) calls the 'ratio equilibria', a subset of the full class of Lindahl equilibria. Ratio equilibria are those in which each individual's tax payment is equal to a personalised scalar (here denoted τ_i for individual i) multiplied by the total cost of public goods, such that all individuals maximise utility and agree on the optimal level of public goods. I thank a referee for prompting this note.

second, the cost of the public goods is fully covered by tax payments. These requirements can be stated formally.

DEFINITION 1. *A first-best Lindahl equilibrium: The policy $\{\{\tau_i^{\text{FB}}\}_i, G^{\text{FB}}\}$ is a first-best Lindahl equilibrium if and only if individuals maximise utility and the following conditions hold:*

$$G_i = G^{\text{FB}} \forall i,$$

$$\sum_{i \in I} \tau_i^{\text{FB}} = 1.$$

Note that this definition describes ‘a’ first-best Lindahl equilibrium, so uniqueness is not implied. With a general form for the ability production function, multiple policies $\{\{\tau_i^{\text{FB}}\}_i, G^{\text{FB}}\}$ may satisfy the conditions of Definition 1 and the analytical results below apply to any first-best Lindahl equilibrium. For example, if public goods spending initially benefits all individuals proportionally but, at higher levels, benefits high-ability individuals more than proportionally, there may exist an optimum with small G and flat tax shares and another optimum with large G and progressive tax shares. Further study of this possibility, including empirical work, is an important next step in the theory and application of CBBT.

Next, I turn to characterising optimal policy with conditions on its components, G^{FB} and $\{\tau_i^{\text{FB}}\}_i$, that depend on only (at least potentially) estimable statistics. All proofs of the following results are collected in the Appendix.

1.2.1. First-best level of public goods spending (Samuelson rule)

First, I can derive the version of the Samuelson rule (1955) that determines the optimal level of public spending in this model. This version turns out to depend on an elasticity that may be a fruitful target for empirical research, defined as follows:

DEFINITION 2. *Define the elasticity of individual i 's income-earning ability with respect to public goods, $\varepsilon_i^G(G)$, as:*

$$\varepsilon_i^G(G) = \frac{f_G(a_i, G)}{f(a_i, G)} G.$$

Using this elasticity, I can state the following result.

PROPOSITION 1. *Samuelson rule: if the policy $\{\{\tau_i^{\text{FB}}\}_i, G^{\text{FB}}\}$ is a first-best Lindahl equilibrium, then the following condition is satisfied:*

$$\sum_{i \in I} \varepsilon_i^G(G^{\text{FB}}) y_i = p G^{\text{FB}}. \quad (7)$$

Equivalently, the sum of individuals' marginal rates of substitution of private consumption for public goods is equal to the marginal rate of transformation between them, when individuals choose according to (4), (5) and (6).

This result simplifies further when the ability production function takes the familiar multiplicative or Cobb-Douglas forms. The key to these simplifications is that the elasticity defined above becomes constant across individuals. In particular:

LEMMA 1. *If the ability production function is multiplicative, such that $f(a_i, G) = h(a_i)g(G)$ for some differentiable functions $h(a_i)$, $g(G)$, both $\mathbb{R}^+ \rightarrow \mathbb{R}^+$, then at a given G the elasticity of individual i 's income-earning ability with respect to public goods, $\varepsilon_i^G(G)$, satisfies:*

$$\varepsilon_i^G(G) = \frac{g'(G)G}{g(G)} = \varepsilon_j^G(G) \equiv \varepsilon^G(G) \quad \forall i, j.$$

Furthermore, if $g(G) = G^\gamma$ for some $\gamma > 0$,

$$\varepsilon^G(G) = \gamma \quad \forall G.$$

These results lead to the following Corollary of Proposition 1.

COROLLARY 1. *If the policy $\{\{\tau_i^{FB}\}_i, G^{FB}\}$ is a first-best Lindahl equilibrium, and if the ability production function is multiplicative, such that $f(a_i, G) = h(a_i)g(G)$ for some differentiable functions $h(a_i)$, $g(G)$, both $\mathbb{R}^+ \rightarrow \mathbb{R}^+$, then G^{FB} satisfies:*

$$\varepsilon^G(G^{FB}) \sum_{i \in I} y_i = pG^{FB}. \tag{8}$$

Furthermore, if $g(G) = g^\gamma$ for some $\gamma > 0$, then spending on G^{FB} is a share γ of total output:

$$\gamma \sum_{i \in I} y_i = pG^{FB}.$$

In words, Corollary 1 says that, in the case of a multiplicative ability production function, society ought to devote a fraction of its output equal to the (uniform) elasticity of income-earning ability with respect to public goods, $\varepsilon^G(G^{FB})$, toward funding the public goods that magnify endowed abilities. If income-earning ability is log-linear in public goods with coefficient γ – such as the familiar Cobb-Douglas form – that fraction is simply γ .

This Corollary provides a neat illustration of how taxes in the Lindahl equilibrium enable private market efficiency in the public sector. Suppose that at the efficient outcome the provider of public goods is paid a price for its output equal to the marginal product of public goods. That marginal product is $\sum f_G(a_i, G)l_i$, where $l_i = y_i/f(a_i, G)$ is the (held fixed) labour effort for individual i . The total payments to public goods would then be $G \sum f_G(a_i, G)y_i/f(a_i, G)$. Now impose Cobb-Douglas and this becomes $\gamma \sum y_i$, so that the total payments to public goods equals the share γ of total income, just as for a factor of production in a competitive equilibrium.

1.2.2. First-best average tax rates

Next, we turn to characterising the taxes paid by each individual in the first-best allocation. As with the optimal level of public goods spending, our results can be expressed in terms of a potentially estimable elasticity.

DEFINITION 3. Define the Hicksian partial elasticity of complementarity between public goods and endowed ability, $\theta_i^{G,a}$, as:

$$\theta_i^{G,a} = \frac{f_{G,a}(a_i, G)f(a_i, G)}{f_G(a_i, G)f_a(a_i, G)}, \quad (9)$$

at a given G .

The Hicksian partial elasticity of complementarity has received attention in recent optimal tax analyses that include human capital (see Bovenberg and Jacobs, 2011; Stantcheva, 2015, and the citations therein). In this article's context, it captures the degree to which public goods and endowed ability magnify each other in determining income-earning ability. If $\theta_i^{G,a} \leq 0$, endowed ability and public goods are not complements in the production of income-earning ability. If $\theta_i^{G,a} \in (0, 1)$ the elasticity of income-earning ability with respect to the level of public goods spending is positive but decreasing in endowed ability; if $\theta_i^{G,a} > 1$, the elasticity of income-earning ability with respect to the level of public goods spending is increasing in endowed ability. We can then state the following result.

PROPOSITION 2. Assume public goods and endowed ability have positive partial effects on income-earning ability: i.e. $f_G(a_i, G) > 0$ and $f_a(a_i, G) > 0$ for all i . If $\{\{\tau_i^{\text{FB}}\}_i, G^{\text{FB}}\}$ is a first-best Lindahl equilibrium, with corresponding incomes $\{y_i^{\text{FB}}\}_i$ then the optimal average tax rate on individual i increases (decreases) with endowed ability iff the Hicksian partial elasticity of complementarity between public goods and endowed ability is greater than (less than) one. Formally, define the optimal average tax rate for individual i as $\text{ATR}_i^{\text{FB}} = \tau_i^{\text{FB}} p G^{\text{FB}} / y_i^{\text{FB}}$. Then,

$$\frac{\partial \text{ATR}_i^{\text{FB}}}{\partial a_i} > (<) 0 \Leftrightarrow \theta_i^{G,a}(G^{\text{FB}}) > (<) 1.$$

The intuition for Proposition 2 is as follows. Public goods are assumed to magnify all abilities.¹⁷ If the effect of public goods on ability has a greater elasticity to endowed ability than does ability itself, that means that there is a complementarity between public goods and ability.¹⁸ Thus, public goods spending is a source of greater benefit for those with greater endowed ability and optimal benefit-based taxation adjusts their taxes up accordingly. Again, note that if the ability production function takes a form in which the effect of G varies in a complicated way with a , multiple equilibria with different degrees of tax progressivity (as well as levels of G) could satisfy the conditions for a first-best Lindahl equilibrium.

As with Proposition 1, I can simplify this result if I assume a multiplicative form for the ability production function:

¹⁷ In principle, this need not be the case, though if I restrict attention to modern developed economies it is difficult to imagine that any individual's income-earning ability is lowered, in an absolute sense, by the existence of the state.

¹⁸ One can define the elasticity of individual i 's income-earning ability with respect to endowed ability as $\varepsilon_i^a(G) = [f_a(a_i, G)/f(a_i, G)]a_i$ and the elasticity of the effect of public goods on individual i 's income-earning ability with respect to endowed ability as: $\varepsilon_i^{G,a}(G) = [f_{G,a}(a_i, G)/f_G(a_i, G)]a_i$. The Hicksian partial elasticity of complementarity is the ratio of the latter to the former.

LEMMA 2. *If the ability production function is multiplicative, such that $f(a_i, G) = h(a_i)g(G)$ for some differentiable functions $h(a_i), g(G)$, both $\mathbb{R}^+ \rightarrow \mathbb{R}^+$, then $\theta_i^{G,a} = 1$.*

The intuition for this Lemma is that the multiplicative production function means the elasticity of income-earning ability with respect to the level of public goods spending is the same no matter the level of endowed ability. These results and Lemma 1 lead to the following corollary of Proposition 2.

COROLLARY 2. *If $\{\{\tau_i^{FB}\}_i, G^{FB}\}$ is a first-best Lindahl equilibrium, and if the ability production function is multiplicative, such that $f(a_i, G) = h(a_i)g(G)$ for some differentiable functions $h(a_i), g(G)$, both $\mathbb{R}^+ \rightarrow \mathbb{R}^+$, then the optimal tax policy is a uniform average tax rate for all $i \in I$:*

$$ATR_i^{FB} = \varepsilon^G(G^{FB}) \quad \forall i. \tag{10}$$

Furthermore, if $g(G) = g^\gamma$ for some $\gamma > 0$, then the optimal tax policy is a uniform tax rate of γ :

$$ATR_i^{FB} = \gamma \quad \forall i. \tag{11}$$

Naturally, this Corollary and Corollary 1 are linked, in that total spending on public goods as a share of output equals this flat average tax rate.

1.3. Relation to Sacrifice Theories

Mill’s (1871 [1994]) principle of equal sacrifice was the theory of optimal taxation that displaced Smith’s CBBT in the intellectual history of optimal tax theory. Under equal sacrifice, tax burdens are to be distributed so that each taxpayer feels the same sacrifice from contributing to public goods.

A relationship between benefit-based taxation and equal sacrifice has long been intuited. For example, in an early critique of the Mirrleesian approach Feldstein (1976) wrote: ‘Nozick (1974) has recently presented an extensive criticism of the use of utilitarian principles to justify the redistribution of income and wealth...In this context, the principle of benefit taxation or of tax schedules that impose equal utility sacrifice have an appeal that is clearly lacking in the utilitarian framework’. (p. 85)

Using the model above, and assuming some specific but familiar functional forms, I can show a quantitative connection that is consistent with this intuition.

To link these principles requires defining ‘sacrifice’. To do so, I consider a hypothetical state in which public goods are provided at their optimal level, G^{FB} , for no cost. Sacrifice for individual i is then defined as the decrease in utility of moving from an undistorted optimal position in that hypothetical state to the planner’s chosen allocation.

In the hypothetical, free-public-goods state, the individual sets:

$$u_c \left[c_i^H, \frac{y_i^H}{f(a_i, G^{FB})} \right] = \frac{-1}{f(a_i, G^{FB})} u_l \left[c_i^H, \frac{y_i^H}{f(a_i, G^{FB})} \right],$$

where G^{FB} satisfies Proposition 1, $c_i = y_i$ (there are no taxes), and the H superscript denotes the ‘hypothetical’ free public goods scenario.

To calculate utility levels, I need to specify the utility functional form. I assume the following familiar specification (which is consistent with Mill’s 1871 [1994] discussion of marginal utility above a minimum income).

$$U(c_i, l_i) = \ln(c_i) - \frac{1}{\sigma} \left[\frac{y_i}{f(a_i, G)} \right]^\sigma, \quad (12)$$

so the individual’s optimality condition in the hypothetical state is simply:

$$c_i^H = y_i^H = f(a_i, G^{\text{FB}}),$$

and utility levels in the hypothetical free public goods state are:

$$U(c_i^H, l_i^H) = \ln(c_i^H) - \frac{1}{\sigma} \left[\frac{y_i^H}{f(a_i, G^{\text{FB}})} \right]^\sigma. \quad (13)$$

Finally, sacrifice for individual i at the first-best Lindahl equilibrium allocation $\{c_i^{\text{FB}}, y_i^{\text{FB}}, G^{\text{FB}}\}$ is:

$$S_i^{\text{FB}} = U(c_i^H, l_i^H) - U(c_i^{\text{FB}}, l_i^{\text{FB}}). \quad (14)$$

With these, I can show the following result.

PROPOSITION 3. *If $\{\{c_i^{\text{FB}}\}_i, G^{\text{FB}}\}$ is a first-best Lindahl equilibrium, the utility function takes the form in (12), and the ability production function is log-linear in public goods, that is, of the form $f(a_i, G) = h(a_i)G^\gamma$ for some differentiable function $h(a_i); \mathbb{R}^+ \rightarrow \mathbb{R}^+$ and $\gamma > 0$, then $S_i^{\text{FB}} = S_j^{\text{FB}}$ for all types i, j : that is, the policy generates equal sacrifice for all individuals.*

It is important to note that this exact equivalence result holds only under a special set of functional form assumptions. Nevertheless, because those assumptions are familiar and arguably plausible, the connection in Proposition 3 hints at more than a superficial relationship. In particular, it suggests that two of the most prominent alternatives to the conventional social welfare maximisation criterion – the CBBT logic of Smith (1776 [1991]) and the equal sacrifice logic of Mill (1871 [1994]) – push optimal policy in a similar direction away from the conventionally-optimal design. Consistent with this finding is the observation that Mill, who rejected benefit-based taxation in favour of equal sacrifice taxation, nevertheless endorsed a similar tax schedule as Smith and invoked Smith as inspiration for his reasoning.¹⁹ Therefore, to the extent that either of these principles presents a plausible alternative – from either a positive or a normative perspective – to the conventional approach, this connection may strengthen the claims of the other.

2. Quantitative Characterisation of Constrained–Optimal (second-best) Classical Benefit-based Taxes

A persistent critique of Lindahlian benefit-based reasoning (Mill, 1871 [1994]; Samuelson, 1954), is that individuals’ benefits are unobservable. How to pursue

¹⁹ Both endorsed proportional taxation above a minimum level of income.

benefit-based policy when constrained by this unobservability has been a topic of intense study for decades (Foley, 1970). When one uses the classical view of benefit-based taxation, however, the problem of unobservable benefit is converted to the problem of unobservable ability – exactly the constraint at the heart of modern Mirrleesian tax theory. In this Section, I take advantage of that feature of CBBT and use familiar methods to study constrained optimal benefit-based taxation. In particular, the purpose of this Section is to understand how constrained-optimal benefit-based taxation responds quantitatively to the way in which public goods and endowed ability interact to produce income-earning ability. I consider illustrative cases of that interaction to show the wide variety of tax policies that may be consistent with classical benefit-based reasoning.

2.1. *Specifying the Constrained Planner's Problem*

Like many alternatives to the standard welfarist objectives in modern optimal tax theory, the classical benefit-based principle does not rank allocations other than its most-preferred. As incentive constraints will force the policy away from its first-best optimum, we need to be able to complete this ranking to identify the second-best policy. The approach taken here applies a simplified version of the technique developed in Weinzierl (2014); see that paper for a more general approach.

To rank the set of feasible and incentive compatible allocations according to the classical benefit-based view, I specify an objective function that penalises symmetric deviations from that principle's first-best allocation.²⁰ I specify a simple version of such an objective function that nevertheless satisfies the Pareto criterion that allocations yielding greater utilities for some and no less utility for all are (at least weakly) preferred, thus avoiding the concerns of Kaplow and Shavell (2001). In particular, the objective function has a kink at its most-preferred allocation, so that the welfare loss due to a deviation below that allocation is greater than the welfare gain due to a similar-sized deviation above it.

Formally, recall the optimal feasible allocation that we derived and characterised in the previous Section $\{c_i^{\text{FB}}, y_i^{\text{FB}}, G^{\text{FB}}\}_i$. The objective function, denoted \mathcal{W} , is defined as follows:²¹

$$\mathcal{W}(\{c_i, y_i, G\}_i, \{c_i^{\text{FB}}, y_i^{\text{FB}}, G^{\text{FB}}\}_i) = \sum_{i=1}^I V \left\{ U \left[c_i, \frac{y_i}{f(a_i, G)} \right], U \left[c_i^{\text{FB}}, \frac{y_i^{\text{FB}}}{f(a_i, G^{\text{FB}})} \right] \right\}, \quad (15)$$

²⁰ As far as I am aware, no previous work has studied how to obtain a complete ranking of allocations based on classical benefit-based reasoning (i.e. where benefits are linked to income-earning abilities). My approach is, therefore, by necessity somewhat speculative and future research may, of course, discover preferable alternative specifications.

²¹ Readers familiar with Weinzierl (2014) will notice a strong similarity between the treatment of benefit-based taxation in this article and equal sacrifice taxation in that paper. Though subsection 1.3 showed that these criteria may yield quantitatively similar results in some cases, the similarity of treatment is not due to this connection but rather to the requirement of an objective function that punishes deviations from an optimal allocation but respects Pareto efficiency.

where:

$$V(U_i, U_i^{FB}) = \begin{cases} \delta(U_i - U_i^{FB}) & \text{if } U_i^{FB} < U_i \\ (U_i - U_i^{FB}) & \text{if } U_i^{FB} \geq U_i \end{cases}, \tag{16}$$

for scalar $\delta : 0 \leq \delta \leq 1$.

The objective function in expressions (15) and (16) applies weight δ , where $0 \leq \delta \leq 1$, to deviations of individual utility above the allocation $\{c_i^{FB}, y_i^{FB}, G^{FB}\}_i$. The asymmetric punishment of downward deviations from the benchmark allocation implied by (the strict case of) $\delta \leq 1$ rejects the utilitarian idea that the distribution of utility across individuals is irrelevant. The assumption that $\delta \geq 0$ respects a weak form of Pareto efficiency ($\delta > 0$ would respect a strong form). Though the non-differentiability caused by $\delta < 1$ is technically inconvenient, it respects the classical benefit-based criterion’s priorities. In contrast, a smooth loss function would imply local indifference to symmetric deviations from the most-preferred allocation, violating the principle that individuals who benefit equally from the activities of the state ought to pay equal taxes.

With this objective function, I can now state a planner’s problem that closely mimics that of the conventional optimal tax model:

PROBLEM 1. *Classical benefit-based planner’s problem:*

$$\max_{\{c_i, y_i, G\}_{i=1}^I \in \{\mathbb{F} \cap \mathbb{IC}\}} \mathcal{W}(\{c_i, y_i, G\}_i, \{c_i^{FB}, y_i^{FB}, G^{FB}\}_i), \tag{17}$$

where \mathcal{W} is defined in expressions (15) and (16), \mathbb{F} denotes the set of feasible allocations for the economy:

$$\mathbb{F} = \left\{ \{c_i, y_i, G\}_{i=1}^I : \sum_{i=1}^I (y_i - c_i) \geq G \right\}, \tag{18}$$

and \mathbb{IC} denotes the set of incentive compatible allocations:

$$\mathbb{IC} = \left\{ \{c_i, y_i, G\}_{i=1}^I : U[c_i, y_i/f(a_i, G)] \geq U[c_j, y_j/f(a_i, G)] \text{ for all } i, j \in \{1, 2, \dots, I\} \right\}. \tag{19}$$

2.2. Parameterisation of the Model, Including the Ability Production Function

I assume a familiar form for the individual utility function, modified to include the effect of public goods on ability:

$$U[c_i, y_i/f(a_i, G)] = \frac{1}{1 - \psi} (c_i)^{1 - \psi} - \frac{1}{\sigma} \left[\frac{y_i}{f(a_i, G)} \right]^\sigma.$$

For this utility function, I assume $\psi = 1.5$ and $\sigma = 3$.

The novel specification required in this article is how income-earning abilities are generated, that is, the form of the function $f(a_i, G)$. It is important to be clear that this article does not base on direct empirical evidence its assumptions on the form of $f(a_i, G)$, though estimating such a relationship may be important for optimal tax

analysis (both in this context and more generally). Instead, I will consider a flexible specification for $f(a_i, G)$ in order to illustrate the potential implications of the model. As will be clear, even a relatively simple but flexible specification is capable of yielding optimal policies that span a wide range of policy patterns (i.e. in terms of progressivity and the role of government). Those numerical findings are consistent with the perpetual uncertainty in benefit-based thinking over whether the rich or poor benefit more from public goods. One way to interpret a contribution of this article is to put that debate into formal terms, clarifying the implicit conditions that resolve that uncertainty and making clear the empirical evidence required to resolve it.

Specifying $f(a_i, G)$ poses a number of challenges: it ought to produce plausible distributions of ability given a realistic value of G ; it ought to allow endowed ability to be related to income-earning ability in a range of ways; and it ought to be as analytically simple and empirically applicable as possible.

To address these challenges, I map individuals' endowed types to fixed percentile positions in an ability distribution that is endogenous to the level of G .²² Specifically, I will assume that income-earning ability is lognormally distributed and that the parameters of that distribution depend on the level of G .²³ Then, for any value of G , individual i will have income-earning ability equal to the inverse cumulative distribution function (cdf) of the resulting income-earning distribution at i 's fixed percentile position. Formally, I can write as in the following:

$$P(a_i) = \Phi \left[\frac{\ln f(a_i, G) - \mu(G)}{\sigma(G)} \right],$$

where Φ is the standard normal cdf, $P(a_i)$ is the percentile for type a_i , and where the mean $\mu(G)$ and standard deviation $\sigma(G)$ of the lognormal income-earning distribution depend on G . Income-earning ability can then be derived, using the inverse cdf and its parameters, so that for type i :

$$\ln f(a_i, G) = \sigma(G)\Phi^{-1}[P(a_i)] + \mu(G).$$

Note that this structure assumes that public goods do not affect the ordering of individuals in the ability distribution. While relaxing that assumption may be of interest, it is left for future work. One reason to assume that $f(a_i, G)$ is order-preserving, at least in the context of a positive optimal tax analysis, is that it may be difficult to convince most people of a specific alternative.

Lacking evidence on the forms of $\mu(G)$ and $\sigma(G)$, I assume the following simple forms:

$$\begin{aligned} \mu(G) &= \bar{\mu} - \frac{\xi}{G} + \ln G^\gamma, \\ \sigma(G) &= \bar{\sigma} + \beta_1 G + \beta_2 G^2, \end{aligned} \tag{20}$$

where ξ , γ , $\bar{\mu}$, and $\bar{\sigma}$ are constants.

²² The equality of opportunity literature, for instance, Roemer (2009), has used a similar fixed-positions technique.

²³ Though the lognormal distribution is relatively simple and, therefore, has the advantage of making clear the impact of G on the distribution of abilities, income distributions are better described by the so-called double-Pareto-lognormal (DPLN) distributions. Future research may fruitfully extend this analysis to that case.

Though simple, the forms assumed in (20) have a few appealing features.

First, this specification implies that all income-earning abilities approach zero as the state disappears (i.e. $\lim_{G \rightarrow 0} \mu(G) = -\infty$ because of the term $-\zeta/G$). A degenerate ability distribution at zero is a natural starting point for analysis, as it addresses the classic critique of (non-classical) benefit-based theory: namely, that without a state income-earning abilities would be negligible (Murphy and Nagel, 2002, p. 17).

Second, the specification lends itself to estimation with (time-series, cross-country, or panel) data on (per capita) government expenditure and the distributional parameters of empirical ability distributions. As the empirical viability of benefit-based taxation has always been considered an obstacle, specifications that might conceivably be taken to the data may be important.

Third, expression (20) implies that, for $\beta_1 = \beta_2 = 0$, the ability production function approaches a familiar form as ζ becomes small. That is:

$$\lim_{\zeta \rightarrow 0} f(a_i, G) = \eta(a_i)G^\gamma, \quad (21)$$

where $\eta(a_i) = \exp\{\bar{\sigma}\Phi^{-1}[P(a_i)] + \bar{\mu}\}$ is constant with respect to G . Expression (21) is the simple multiplicative form for which Corollaries 1 and 2 were derived in the previous Section. In fact, this expression is log-linear in G , for which I derived the especially clean results that spending on public goods as a share of total output and the flat average tax rate are equal to γ . It may be interesting, and reassuring, to note the (informal) intuition of Smith (1776 [1991]) was that a classical benefit-based logic would lead to just such a proportional tax.²⁴

To demonstrate the range of potential optimal policies under CBBT, I construct three sets of parameter values that are consistent with the empirical status quo but that have dramatically different optimal policy implications. Specifically, I choose three sets of values of $\bar{\mu}$, ζ , γ , $\bar{\sigma}$, β_1 and β_2 so that the simulated economy in each case – assuming the status quo tax policy – will be a close match to the overall share of spending on public goods in total income and the mean and variance of a lognormal distribution fitted to the US wage distribution.²⁵ For status quo tax policy, I use an approximation of the current US tax system with a flat tax rate of 40% and a lump sum grant equal to the excess of tax revenue over government spending. To calibrate the target ability distribution roughly, I use data on reported annual earnings and hours (as well as sample weights) from the National Longitudinal Survey of Youth 1996 survey to calculate the distribution of wages, a standard (though imperfect) empirical proxy for ability. The lognormal distribution that provides the best fit to these calculated wages has a mean of 2.5 and a standard deviation of 0.7. Finally, US national accounts show federal government spending to be approximately 7.6% of GDP.

²⁴ Note that Smith does endorse progressive taxation more generally, largely out of concern for the very poor. This mixed perspective is exactly that generated by the positive optimal tax analysis in this article, where a conventional utilitarian logic motivates redistribution but a classical benefits-based logic limits the extent of that redistribution and its associated progressivity. See also the earlier footnote on Smith's mixed normative reasoning.

²⁵ The simulation has a distribution of individuals (indexed by their percentile position in the endowed ability distribution) choose labour effort to maximise utility, taking into account the ability production function, flat tax rate, and share of G in output.

Table 1 shows these three sets of parameter values. The first column shows what I call the ‘baseline’ specification, where $\beta_1 = \beta_2 = 0$ and thus $f(a_i, G)$ satisfies expression (21). The remaining columns of Table 1 show variations on the baseline scenario in bold. The second column of Table 1 shows an ‘expansive state’ specification in which the baseline value of $\bar{\mu}$ is decreased to allow for a larger γ . In this scenario, public goods have greater power to magnify endowed ability. In the third column, I return to the baseline values of $\bar{\mu}$, choose an intermediate value of γ and relax the baseline assumption that β_1 and β_2 are zero, instead considering a case in which increases in G are relatively more beneficial for workers with high endowed ability. Formally, to the extent that increases in G raise $\sigma(G)$, those higher in the endowed ability distribution will benefit more from public goods. I label this scenario the ‘progressive’ case, where $\beta_1 > 0$ is partially offset by $\beta_2 < 0$, so that a finite (and reasonable) value for G is chosen and $\bar{\sigma}$ is smaller.

While all three sets of parameter values are consistent with the same size of government and shape of the ability distribution under an approximation of the status quo US tax policy, they imply quite different constrained-optimal policies. To obtain those policies, we simulate the planner’s problem in expressions (17), (18) and (19) with $\delta = 0$, corresponding to a strict version of CBBT.

In Figure 1, we plot average tax rates along the ability distribution in the second-best optimal benefit-based policy for the three scenarios. The message from these simulations is that optimal benefit-based taxation in its classical form can support a variety of sizes for government activity and a range of degrees of progressivity. That is, the shares of G in total income in these three optimal policies are 7.6%, 15.2%, and 11.8%, and the last of the three exhibits substantial progressivity. Benefit-based taxation is not necessarily ‘libertarian’ in the sense of a minimal state; nor is it inconsistent with progressive taxation.²⁶

Table 1
Parameterisations of the Ability Production Function

Parameter	1 (Baseline)	2 (Expansive state)	3 (Progressive)
$\bar{\mu}$	2.36	2.20	2.36
ξ	0.01	0.01	0.01
γ	0.076	0.152	0.114
$\bar{\sigma}$	0.70	0.70	0.30
β_1	0.00	0.00	0.06
β_2	0.00	0.00	-0.001
<i>Empirical targets under Status Quo tax policy</i>			
$\mu(G) = 2.5$	2.5	2.5	2.6
$\sigma(G) = 0.7$	0.7	0.7	0.7
$G / \sum_i y_i$	0.076	0.076	0.076

²⁶ Both of these features are, in fact, consistent with Smith’s view of taxation. Smith (1776 [1991]) writes, ‘The third and last duty of the sovereign or commonwealth is that of erecting and maintaining those public institutions and those public works...necessary for the defence of the society, and for the administration of justice . . . for facilitating the commerce of the society, and those for promoting the instruction of the people’ (p. 473). Also see the earlier footnote on Smith’s endorsement of progressivity.

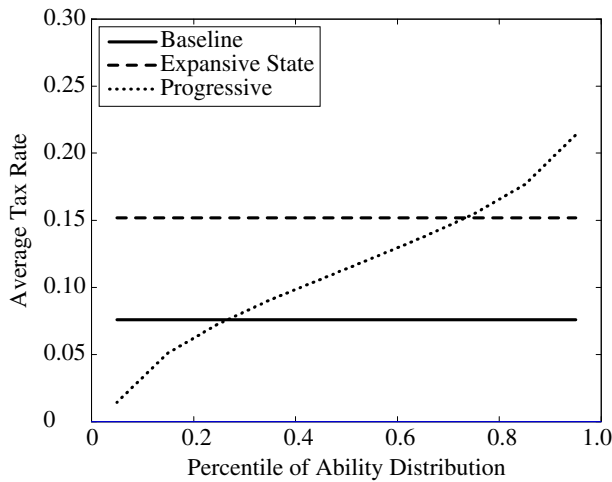


Fig. 1. Average Tax Rates in the Constrained-optimal Policy Under Three Sets of Parameters from Table 1 Note. Colour figure can be viewed at wileyonlinelibrary.com.

This flexibility is arguably behind CBBT's long-standing and apparently widespread appeal. As I show in a companion paper (Weinzierl, 2016), support for CBBT among the public appears to be widespread and to cut across political and demographic categories. Rather than a niche principle to which a small group of iconoclasts subscribe, CBBT is a unifying idea that can be consistent with a range of views on both the impact of public goods spending on the ability distribution and, therefore, optimal government tax and expenditure policy more generally.²⁷

That said, Figure 1 also shows that CBBT has sharp enough implications to inform the debate over optimal policy. First, disagreement over the incidence of benefits has a long history at the heart of the analysis of benefit-based taxation. The formalisation of CBBT in this article provides a way to discipline that debate and describe one's position in a more transparent, testable form. Second, optimal taxes under CBBT include negative average tax rates only if individuals are harmed by the activities of the state, an unlikely situation in modern developed economies. As a consequence, supporters of greater redistribution are less likely to support CBBT, and this feature of CBBT makes it easier for voters and policy makers to understand the positions they defend. As discussed at length in Section 3, evidence suggests that a mixture of a non-redistributive principle such as CBBT with a redistributive principle such as utilitarianism is likely to provide the best description of the overall criterion that prevails in public judgments of tax policy. Understanding the implications of CBBT, both in general and given a specific ability production function, can help individuals decide what weight they put on the components of that mixture.

²⁷ To take a specific example, both Barack Obama and Mitt Romney – who as shown in the Appendix express quite different views of the role of the state and optimal progressivity – invoked reasoning related to CBBT as justification for their positions in the 2012 US presidential election.

Table 2
Hicksian Partial Elasticities of Complementarity $\theta_i^{G,a}$

Percentile of ability distribution	1 (Baseline)	2 (Expansive state)	3 (Progressive)
10	1.00	1.00	5.91
30	1.00	1.00	1.82
50	1.00	1.00	1.55
70	1.00	1.00	1.43
90	1.00	1.00	1.40

We can connect the results shown in Figure 1 to the analysis of first-best policy from Section 2 by calculating the realised values of the Hicksian partial elasticity of complementarity between public goods and endowed ability, $\theta_i^{G,a}$ as defined in expression (9). In Table 2 we show these values for the three cases plotted in Figure 1.

As expected, the multiplicative form for the ability production function in the first two cases yields $\theta_i^{G,a} = 1.00$ for all types, while case 3 shows that the progressivity of taxes is connected to the strong complementarity assumed between public goods and endowed ability in that case. To the extent that the Hicksian partial elasticity is a potential object of empirical study, it may therefore provide an important target for calibration exercises using this model.

3. Quantitative Results Under a Mix of the Classical Benefit-based and Utilitarian Criteria for Optimal Taxes

While the benefit-based criterion plays a prominent role in public rhetoric over taxation, it is unlikely to be the only principle that guides the design of taxes. In particular, the utilitarian criterion that dominates quantitative work in the optimal tax literature almost certainly plays a role – perhaps a very large role – in policy evaluation. The ubiquity of redistribution in advanced economies requires, given that benefit-based taxation is unlikely to recommend negative average tax rates on any individual, that some other principle exerts influence on policy; utilitarianism is a natural candidate for that alternative. In fact, the utilitarian logic is consistent with non-benefit-based arguments with which American politicians and analysts justify tax policy preferences, for instance when stressing the importance of helping the ‘needy’ or those who are ‘struggling’ to make necessary purchases.²⁸

In this context it may be of interest that Richard Musgrave (p. 46) (Buchanan and Musgrave, 1999), toward the end of an illustrious career as a scholar of taxation and the role of the state, wrote: ‘Moreover, observers such as myself who tend to be egalitarian should not rule out the norm of Lockean entitlement to earnings ... as an

²⁸ At the same time, redistribution – a key source of welfare gains in the standard approach – is often explicitly rejected as a justification for tax policy, even by policy makers who might be expected to support it. For example, the statement by President Obama given in the Introduction includes the claim that ‘everybody pays’. Also consider a statement he made in 2012: ‘So when we have debates now about our tax policy, when we have debates now about the Buffett rule that we’ve been talking about, where we say if you make a million dollars a year or more you shouldn’t pay a lower tax rate than your secretary, that is not an argument about redistribution. That is an argument about growth’ (Obama, 2012).

alternative criterion that deserves consideration. Most people, I suggest, would wish to assign some weight to both norms ... I also think that entitlement to earnings, the Lockean and Adam Smith tradition, has its merit. I would give it, say, one-quarter weight with three-quarters to the Rawlsian concept'.

In this Section, I demonstrate that an objective for taxation that incorporates both the classical benefit-based and utilitarian criteria can, as this discussion suggests, yield optimal policies consistent with important features of existing policy. The first step in that demonstration is to show how such a mixed objective can be formalised.

3.1. Formalising a Mixed Objective Function

To combine CBBT and utilitarianism into one objective, I take advantage of a convenient feature of the formalisation of the classical benefit-based planner's problem in expressions (17), (18) and (19). Recall that, in that problem, the planner's desire to stay near the first-best benefit-based allocation is constrained not only by feasibility and incentive compatibility but also by Pareto efficiency. As discussed earlier, the Pareto efficiency requirement is included by having the planner's objective not only punish deviations of individuals' utilities below their benchmark levels but also reward (at least weakly) deviations of individuals' utilities above those levels. Formally, in expression (16), I assume that $\delta \geq 0$.

The extreme values for δ imply an objective based on only one principle. In the previous Section, for example, I simulated constrained-optimal benefit-based taxation by assuming $\delta = 0$, implying a strict commitment on the part of the planner to the classical benefit-based criterion. At the other extreme, when $\delta = 1$, the planner's objective is equivalent to the conventional utilitarian objective to maximise the unweighted sum of individual utilities. To see this, impose $\delta = 1$ on expression (16) to obtain $V(U_i^{\text{FB}}, U_i) = U_i^{\text{FB}} - U_i$, so that the objective function in expression (15) simplifies to $\mathcal{W}(\cdot) = \sum (U_i - U_i^{\text{FB}})$, and the planner's objective (17) becomes $\max_{\{c_i, y_i, G\}_{i=1}^I \in \{\mathbb{F} \cap \mathbb{IC}\}} \sum (U_i - U_i^{\text{FB}})$. Because the allocations U_i^{FB} do not depend on the chosen allocation $\{c_i, y_i, G\}_{i=1}^I$, this objective is equivalent to: $\max_{\{c_i, y_i, G\}_{i=1}^I \in \{\mathbb{F} \cap \mathbb{IC}\}} \sum U_i$, the utilitarian planner's objective.

Intermediate values of the parameter δ provide, therefore, a simple way to form an objective function that lies between – in a formal sense – the classical benefit-based and utilitarian principles.

3.2. Features of Optimal Policy Under a Mixed Objective

We now show, briefly, how varying δ can yield a wide range of optimal policy results. Most important, we show that optimal policy computed with certain values for δ can match several arguably important features of existing policy that are difficult to reconcile in the conventional model (where $\delta = 1$).

3.2.1. Progressivity

First, I consider optimal average tax rate schedules for a calibrated US ability distribution. I discretise (into 50 types) the ability distribution obtained as the 'baseline' case from Table 1 for the values of $\bar{\mu}$, ζ , γ , $\bar{\sigma}$, β_1 and β_2 : this was the set of parameters that

yielded a small government with a flat average tax rate as constrained-optimal in the previous Section. Figure 2 shows the average tax rate results for four values of δ .

As this Figure shows, the extent of progressivity in the optimal policy is reduced when the classical benefit-based criterion receives greater weight in the objective function (under the baseline specification of the ability production function).

The average tax rate paid by high earners in the conventional utilitarian policy shown in Figure 2 is 46%, while the Congressional Budget Office reports that the average federal tax rate (including all federal taxes) averaged 30% over 2000 to 2010. Though the gap between those figures has many possible explanations, note that the average tax rates assessed under the mixed cases (2 and 3) in Figure 2 are 31% and 39%, substantially closer to what prevails in current US policy.

As for low earners, the CBBT and utilitarian criteria yield sharply conflicting recommendations, with negative average tax rates implied for 0%, 31%, 47% and 57% of the population in the four cases shown in Figure 2. The average tax rates at the 10th percentile of the simulated distribution are approximately 7%, -40%, -80% and -140%, while in reality the CBO reports that in 2011 the lowest earning quintile of Americans received net income (after taxes and transfers) 55% greater than their market income (before taxes and transfers). As with the comparison of taxes on high earners, these rough comparisons suggest that the mixed cases (2 and 3) in Figure 2 provide a relatively good fit with existing US tax policy.

3.2.2. Rank reversals in utility

Second, I consider the distribution of utility levels under the optimal policy in a first-best setting. As has been known since Mirrlees (1971), the first-best utilitarian tax policy recommends that income-earning ability be inversely related to utility levels. That is, consumption is equalised across types, but those with higher income-earning ability are required to exert more labour effort. While the second-best optimal policy

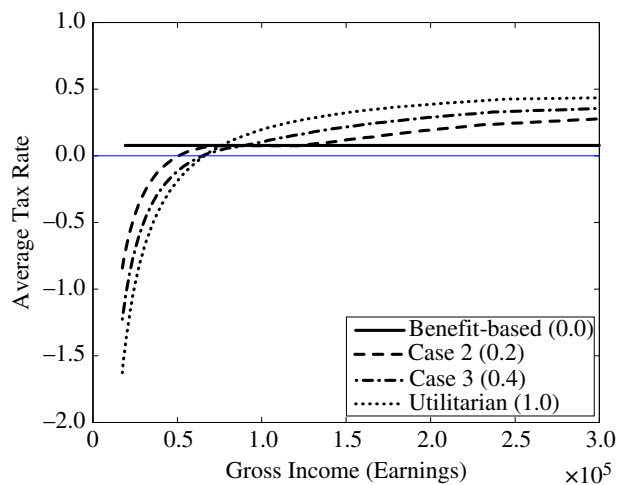


Fig. 2. Average Tax Rates for Four Values of δ (in parentheses), Assuming the Baseline Ability Production Function

Note. Colour figure can be viewed at wileyonlinelibrary.com.

cannot achieve such ‘rank reversals’, many commentators have argued that prevailing norms about economic justice would reject the idea that an unconstrained optimal policy would include them (King, 1983; Saez and Stantcheva, 2015).

The concern over rank reversals relates directly to the ‘endowment taxation’ nature of the conventional modern approach to optimal taxation. Philosophers and legal scholars have extensively analysed this issue (for instances of the latter, see Rakowski, 2000; Shaviro, 2002; Markovits, 2003; Stark, 2005; Zelenak, 2006; Hasen, 2007; Sugin, 2010). Rawls (2001, p. 158) himself worried that an endowment tax ‘would force the more able into those occupations in which earnings were high enough for them to pay off the tax in the required period of time; it would interfere with their liberty to conduct their life within the scope of the principles of justice’.²⁹ The wide-ranging affinity for this critique is apparent when Rawls’ warning is coupled with Nozick’s (1974, p. 169) claim that ‘taxation of earnings from labor is on a par with forced labor’ because ‘it is like forcing the person to work n hours for another’s purpose’.³⁰

To examine the effect of using a mixed objective on the appeal of rank reversals, Figure 3 shows the utility levels achieved by all ability types under the first-best policies in the same four cases as were used to produce Figure 2.

As this Figure shows, the pure utilitarian objective puts in place substantial rank reversals, while the strict benefit-based case has utility positively related to ability in the first-best. Intermediate cases temper the rank reversals of the utilitarian approach.

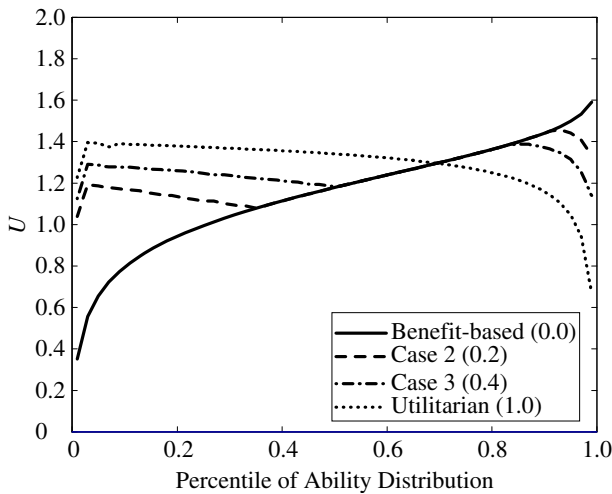


Fig. 3. *Utility Levels by Ability Type in the First-best (full information) Optimal Allocations for Four Values of δ (shown in parentheses), Assuming the Baseline Ability Production Function*
Note. Colour figure can be viewed at wileyonlinelibrary.com.

²⁹ Also see Arneson (2000), who writes: ‘It is better to regard Rawls as making the point that ... it is a flaw that utilitarianism would have the decision about what should be done vary only with the utility total that different acts could achieve’.

³⁰ Stark (2005) offers a detailed argument that the concerns of Rawls (1971) and Nozick (1974) are closely connected.

3.2.3. *Tagging*

Finally, I consider the implications of a mixed criterion for the taxation of personal characteristics, otherwise known as tagging (following Akerlof, 1978). In Mankiw and Weinzierl (2009) and Weinzierl (2014), it is shown that a conventional utilitarian optimal policy would levy substantially greater tax rates on some demographic groups in the US relative to others. In fact, in the modern theory of optimal taxation, tagging is a free lunch and a wide variety of candidate tags exist. Any observable and largely inelastic characteristic across which the distribution of abilities differs ought to affect tax schedules. These conclusions stand in stark contrast to the limited extent of tagging in actual policy. The few examples of existing tagging bear little resemblance to the broad and nuanced application recommended by modern optimal tax theory.

In Weinzierl (2014), a mixed normative criterion that blends utilitarianism with some (empirically-grounded) weight on Mill's principle of equal sacrifice is shown to diminish optimal tagging of these characteristics dramatically and – crucially – much more than it diminishes optimal redistribution.

In this subsection, I find a similar result when the objective function blends utilitarianism with the principle of benefit-based taxation and the effects of public goods on ability are assumed to be independent of the tagged characteristics. The commonality of results across this article and Weinzierl (2014) is related to the connection shown between CBBT and equal sacrifice above.

To show this result, I assume that three tagged groups have different values for the parameters $\mu(G)$ and $\sigma(G)$ in the ability production function (20). The ability distributions for these three groups are shown in Figure 4.

Based on the distributions in Figure 4, a utilitarian criterion will recommend tagging: for example, taxing a member of the 'High' group more than a member of the 'Low' group who has the same income-earning ability. On the other hand, CBBT will not endorse tagging, because individuals with the same ability levels have all benefited equally from the activities of the state, regardless of their tagged group (by assumption).

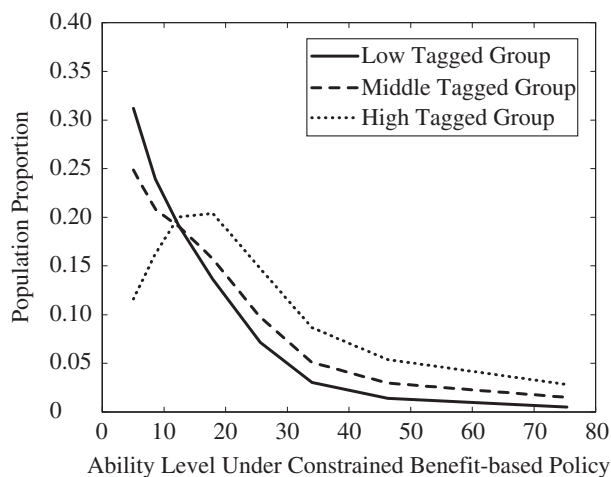


Fig. 4. *Ability Distributions for Three Tagged Groups*

Table 3
Extent of Tagging (Extra tax or transfer rate, in percentages)

	δ	Tagged group			Max ATR
		Low	Middle	High	
Benefit-based	0	0.1	0.0	-0.1	8.2
Case 2	0.20	-0.2	0.0	0.2	33.2
Case 3	0.40	-2.5	0.4	0.6	42.6
Utilitarian	1.00	-9.8	-2.0	7.8	54.3

By varying the value of δ , we can compare the extent of tagging pursued by planners with different weights on the classical benefit-based criterion relative to conventional utilitarianism. I measure the extent of tagging by calculating the total extra taxes paid (or transfers received) by a tagged group, as a share of that group's total income, when a planner initially not allowed to tag is given the power to do so (i.e. to make a person's tax payments depend on which tagged group they belong to). Table 3 shows these tagging results for the same set of values of δ in (16) as were used to generate Figures 2 and 3. It also shows the maximum average tax rate assessed in each simulation.

As this Table shows, the extent of tagging falls dramatically in case 3 and especially case 2 relative to the utilitarian benchmark (case 4), while the progressivity of the tax schedule diminishes much less. In words, these results show that an objective for optimal policy that gives weight to the classical benefit-based criterion as well as utilitarianism can reconcile prevailing policy's limited use of tagging and substantial progressivity.

It is important to clarify that the opposition to tagging under the classical benefit-based criterion depends on the assumption that the effects of public goods are independent of the tagged characteristic. If they are not – for example, if some groups are advantaged more than others by biased public goods provision – then classical benefit-based taxation would recommend taxing based on those characteristics (i.e. tagging). This subtlety is consistent with policies – such as affirmative action in the United States – that are designed to offset the disparities in how endowed ability is translated into income-earning ability across demographic groups.

4. Conclusion

In this article, I have explored whether we might incorporate into modern tax theory a logic for tax design that appears to play a role in public reasoning but that has been largely set aside by theorists. The idea that an individual ought to pay taxes based on the benefit he or she derives from the public goods the government produces has been, for many centuries, an intuitively compelling one. I marry that intuition to two ideas: first, Adam Smith's 'classical' view that the best measure of that benefit is a person's income-earning ability; and second, the plausible assumption that an individual's income-earning ability is a function of both endowed ability and public goods.

As a result, the classical benefit-based taxation turns out to fit seamlessly into a generalised version of the standard Mirrleesian model. I apply Lindahl's approach to assigning taxes based on marginal benefit in that model and show that we can use

familiar modern methods to characterise first-best policy analytically and second-best policy quantitatively. Moreover, the required components of the model are, at least in principle, amenable to empirical analysis using standard data and straightforward elasticities.

Simulations of optimal policy in this model, using an objective function that gives weight to both the conventional utilitarian criterion and the classical benefit-based view, can yield policies that match well several features of existing policy that are difficult to reconcile under the conventional objective. To the extent that such a mixed criterion is a good approximation of prevailing objectives for tax policy, this model may offer a useful approach to a theory of positive optimal taxation.

Harvard Business School

Additional Supporting Information may be found in the online version of this article:

Appendix A. Additional Examples of Classical Benefit-based Rhetoric and Derivations of Results in the Main Text.

Data S1.

References

- Aaron, H. and McGuire, M. (1970). 'Public goods and income distribution', *Econometrica*, vol. 38(Nov. 1), pp. 907–20.
- Aaron, H. and McGuire, M. (1976). 'Reply to Geoffrey Brennan', *Econometrica*, vol. 44(Mar. 1), pp. 401–4.
- Akerlof, G.A. (1978). 'The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning', *The American Economic Review*, vol. 68(1), pp. 8–19.
- Arneson, R.J. (2000). 'Rawls versus utilitarianism in the light of political liberalism', *The Idea of a Political Liberalism: Essays on Rawls*, Lanham MD: Rowman and Littlefield.
- Atkinson, A. and Stiglitz, J. (1980). *Lectures in Public Economics*, New York: McGraw–Hill.
- Best, M. and Kleven, H. (2013). 'Optimal income taxation with career effects of work effort', Working Paper.
- Blakey, R.G. (1914). 'The new income tax', *American Economic Review*, vol. 4(1), pp. 25–46.
- Boadway, R. (2012). *From Optimal Tax Theory to Tax Policy*, Cambridge, MA: MIT Press.
- Bovenberg, L. and Jacobs, B. (2011). 'Optimal taxation of human capital and the earnings function', *Journal of Public Economic Theory*, vol. 13(6), pp. 957–71.
- Brennan, H.G. (1976a). 'The distributional implications of public goods', *Econometrica*, vol. 44(Mar. 1), pp. 391–9.
- Brennan, H.G. (1976b). 'A rejoinder to the Aaron–McGuire reply', *Econometrica*, vol. 44(Mar. 1), pp. 405–7.
- Buchanan, J. and Musgrave, R. (1999). *Public Finance and Public Choice: Two Contrasting Visions of the State*, Cambridge: MIT Press.
- Feldstein, M. (1976). 'On the theory of tax reform', *Journal of Public Economics*, vol. 6(1–2), pp. 77–104.
- Fleurbaey, M. and Maniquet, F. (2006). 'Fair income tax', *Review of Economic Studies*, vol. 73(1), pp. 55–83.
- Foley, D. (1970). 'Lindahl's solution and the core of an economy with public goods', *Econometrica*, vol. 38 (Jan. 1), pp. 66–72.
- Frankfurter, F. and MacLeish, A. (1939). *Law and Politics: Occasional Papers of Felix Frankfurter, 1913–1938* (A. MacLeish and E. F. Prichard Jr, eds.), New York: Harcourt, Brace and Company.
- Gaertner, W. and Schokkaert, E. (2012). *Empirical Social Choice: Questionnaire–Experimental Studies on Distributive Justice*, Cambridge: Cambridge University Press.
- Gelber, A. and Weinzierl, M. (2016). 'Optimal taxation when children's abilities depend on parents' resources', *National Tax Journal*, vol. 69(1), pp. 11–40.
- Hasen, D.M. (2007). 'Liberalism and ability taxation', *Texas Law Review* vol. 85(5), pp. 1057–113.
- Hellwig, M. (2005). 'A utilitarian approach to the provision and pricing of excludable public goods', *Journal of Public Economics*, vol. 89(11–12), pp. 1981–2003.
- Hines, J.R. (2000). 'What is benefit taxation', *Journal of Public Economics*, vol. 75(3), pp. 483–92.

- Institute of Fiscal Studies, Mirrlees, J.A., Adam, S., Besley, T., Blundell, R., Bon, S., Chote, R., Gammie, M., Johnson, P., Myles, G. and Poterba, J. (eds.) (2010). *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press.
- Kaneko, M. (1977a). 'The ratio equilibria and a voting game in a public goods economy', *Journal of Economic Theory*, vol. 16(1), pp. 123–36.
- Kaneko, M. (1977b). 'The ratio equilibria and the core of the voting game $G(N, W)$ in a public goods economy', *Econometrica*, vol. 45(7), pp. 1589–94.
- Kapicka, M. (2006). 'Optimal income taxation with human capital accumulation and limited recordkeeping', *Review of Economic Dynamics*, vol. 9(4), pp. 612–39.
- Kaplow, L. (2008). *Theory of Taxation and Public Economics*, Princeton, NJ: Princeton University Press.
- Kaplow, L. and Shavell, S. (2001). 'Any non-welfarist method of policy assessment violates the Pareto principle', *Journal of Political Economy*, vol. 109(2), pp. 281–6.
- King, M.A. (1983). 'An index of inequality: with applications to horizontal equity and social mobility', *Econometrica*, vol. 51(1), pp. 99–115.
- Lindahl, E. (1919). 'Just taxation – a positive solution' (E. Henderson, transl.), in (R.A. Musgrave and A.T. Peacock, eds.), *Classics in the Theory of Public Finance*, pp. 168–76, Macmillan.
- Lockwood, B.B. and Weinzierl, M.C. (2015). 'De Gustibus non est Taxandum', *Journal of Public Economics*, vol. 124, pp. 74–80.
- Lockwood, B.B. and Weinzierl, M.C. (2016). 'Positive and normative judgments implicit in US tax policy, and the costs of unequal growth and recessions', *Journal of Monetary Economics*, vol. 77, pp. 30–47.
- Mankiw, N.G. and Weinzierl, M.C. (2009). 'The optimal taxation of height', *American Economic Journal: Economic Policy*, vol. 2(1), pp. 155–76.
- Markovits, D. (2003). 'How much redistribution should there be?', Yale Faculty Scholarship Series. Paper No. 412
- Matsumoto, M. (2001). 'Public input provision in an optimal income tax model', *FinanzArchiv*, vol. 58(1), pp. 12–30.
- Mill, J.S. (1871 [1994]). *Principles of Political Economy*, Oxford University Press.
- Mirrlees, J.A. (1971). 'An exploration in the theory of optimal income taxation', *Review of Economic Studies*, vol. 38(2), pp. 175–208.
- Moulin, H. (1987). 'Egalitarian–equivalent cost sharing of a public good', *Econometrica*, vol. 55(3), pp. 963–76.
- Murphy, L. and Nagel, T. (2002). *The Myth of Ownership*, Oxford: Oxford University Press.
- Musgrave, R.A. (1959). *The Theory of Public Finance*, McGraw-Hill.
- Nozick, R. (1974). *Anarchy, State, and Utopia*, New York: Basic Books.
- Obama, B. (2011). 'Remarks at George Washington University', available at: <http://www.presidency.ucsb.edu/ws/index.php?pid=90246&st=benefited+most+from+our+way+of+life&st1=> (last accessed 13 February 2017).
- Obama, B. (2012). 'The President's News Conference With President Juan Manuel Santos Calderon of Colombia in Cartagena', available at: <http://www.presidency.ucsb.edu/ws/index.php?pid=100494&st=so+when+we+have+debates+now+about+our+tax+policy&st1=> (last accessed 13 February 2017).
- Rakowski, E. (2000). 'Can wealth taxes be justified', *NYU Tax Law Review*, vol. 53, pp. 263ff.
- Rawls, J. (1971). *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Rawls, J. (2001). *Justice as Fairness*, Cambridge, MA: Harvard University Press.
- Roemer, J.E. (2009). *Equality of Opportunity*, Cambridge, MA: Harvard University Press.
- Roosevelt, F.D. (1935). 'Message to Congress on Tax Revision', available at: <http://www.presidency.ucsb.edu/ws/index.php?pid=15088&st=&st1=> (last accessed: 13 February 2017).
- Saez, E. and Stantcheva, S. (2015). 'Generalized social welfare weights for optimal tax theory', *American Economic Review*, vol. 106(1), pp. 24–45.
- Salanié, B. (2011). *The Economics of Taxation*, Cambridge, MA: MIT Press.
- Samuelson, P.A. (1954). 'The pure theory of public expenditure', *Review of Economic and Statistics*, vol. (Nov 1) pp. 387–9.
- Samuelson, P.A. (1955). 'Diagrammatic exposition of a theory of public expenditure', *Review of Economics and Statistics*, vol. 37(4), pp. 159–71.
- Seligman, E.R.A. (1908). 'Progressive taxation in theory and practice', *American Economic Association Quarterly*, vol. 9(1 and 2).
- Shapiro, D. (2002). 'Endowment and inequality', in (J.J. Thorndike and D.J. Ventry Jr., eds.), *Tax Justice: The Ongoing Debate*, pp. 123–49, Urban Institute.
- Shapiro, D. (2013). 'The forgotten Henry Simons', NYU Law School Working Paper.
- Sheffrin, S.M. (2013). *Tax fairness and Folk Justice*, New York: Cambridge University Press.
- Simons, H. (1938). *Personal Income Taxation*, Chicago, IL: University of Chicago.
- Smith, A. (1776 [1991]). *An Inquiry into the Nature and Causes of the Wealth of Nations*, New York: Prometheus.
- Stantcheva, S. (2015). 'Optimal taxation and human capital policies over the life cycle', NBER Working Paper 21207.

- Stark, K.J. (2005). 'Enslaving the beachcomber: some thoughts on the liberty objections to endowment taxation', UCLA Law and Economics Research Series.
- Stiglitz, J. (1987). 'Pareto efficient and optimal taxation and the new new welfare economics', in (A.J. Auerbach and M.S. Feldstein, eds.), *Handbook of Public Economics*, Chapter 15, pp. 991–1042, Elsevier.
- Sugin, L. (2010). 'A philosophical objection to the optimal tax model', *NYU Tax Law Review*, vol. 61, pp. 299ff.
- Weinzierl, M. (2014). 'The promise of positive optimal taxation: normative diversity and a role for equal sacrifice', *Journal of Public Economics*, vol. 118(Oct. 1), pp. 128–42.
- Weinzierl, M. (2016). 'Popular acceptance of inequality due to Brute Luck and support for classical benefit-based taxation', NBER Working Paper No. 22462.
- Werning, I. (2007). 'Pareto efficient income taxation', Working Paper, MIT, April.
- West, E.G. and Staaf, R.J. (1979), 'The distributional implications of public goods revisited', *Econometrica*, vol. 47(July 1), v1031–1037.
- Zelenak, L. (2006). 'Taxing endowment', *Duke Law Journal*, vol. 55(6), pp. 1145–81.