NEW YORK UNIVERSITY SCHOOL OF LAW

SPRING 2014

COLLOQUIUM ON TAX POLICY
AND PUBLIC FINANCE

"Revisiting the Classical View of Benefit-based Taxation"

Matthew C. Weinzierl
Harvard Business School

March 25, 2014
NYU Law School
Vanderbilt Hall-208
Time: 4:00-6:00pm
Number 8

SCHEDULE FOR 2014 NYU TAX POLICY COLLOQUIUM
(All sessions meet Thursday 4:00-5:50 p.m., Vanderbilt-208, NYU Law School)

1. January 21 – Saul Levmore, University of Chicago Law School, "From Helmets to Savings and Inheritance Taxes: Regulatory Intensity, Information Revelation, and Internalities." (Main discussion paper); and "Internality Regulation Through Public Choice." (Background paper).

2. January 28 – Fadi Shaheen, Rutgers-School of Law, Newark, "The GAAP Lock-Out Effect and the Investment Behavior of Multinational Firms."; "Evaluating Investments of Locked-Out Earnings (an Outline).

3. February 4 – Nancy Staudt, University of Southern California, Gould School of Law "The Supercharged IPO."

4. February 11 – Thomas J. Brennan, Northwestern University School of Law, "Smooth Retirement Accounts."

5. February 25 – Chris Sanchirico, University of Pennsylvania Law School. "As American as Apple Inc., International Tax and Ownership Nationality."

6. March 4 – James R. Hines, Jr. and Kyle D. Logue, University of Michigan Law School, "Delegating Tax."

7. March 11 – Stephanie Sikes, Wharton School, Accounting Department, University of Pennsylvania. "Cross-Country Evidence on the Relation between Capital Gains Taxes, Risk, and Expected Return."

8. **March 25 – Matthew C. Weinzierl, Harvard Business School, "Revisiting the Classical View of Benefit-based Taxation."**

9. April 1 – Andrew Biggs, American Enterprise Institute, "The Risk to State and Local Budgets Posed by Public Employee Pensions."

10. April 8 – Susannah Camic Tahk, University of Wisconsin Law School, "The Tax War on Poverty"

11. April 15 – Nirupama Rao, NYU Wagner School, "The Price of Liquor is Too Damn High: State Facilitated Collusion and the Implications for Taxes."

12. April 22 – Kimberly Clausing, Reed College, Economics Department, "Lessons for International Tax Reform from the U.S. State Experience under Formulary Apportionment."

13. April 29 – David Gamage, Berkeley Law School, "On Double-Distortion Arguments, Distribution Policy, and the Optimal Choice of Tax Instruments."

14. May 6 – Mitchell Kane, NYU School of Law, "Reflections on the Coherence of Source Rules in International Taxation."

Revisiting the Classical View of Benefit-Based Taxation

Matthew Weinzierl*

Preliminary Draft as of March 13, 2014

Abstract

This paper explores how the persistently popular "classical" logic of benefit based taxation, in which an individual's benefit from public goods is tied to his or her income-earning ability, can be incorporated into modern optimal tax theory. If Lindahl's methods are applied to that view of benefits, first-best optimal policy can be characterized analytically as depending on a few potentially estimable statistics, in particular the coefficient of complementarity between public goods and innate talent. Constrained optimal policy with a Pareto-efficient objective that strikes a balance–controlled by a single parameter–between this principle and the familiar utilitarian criterion can be simulated using conventional constraints and methods. A wide range of optimal policy outcomes can result, including those consistent with existing policies. To the extent that such an objective reflects the mixed normative reasoning behind prevailing policies, this model may offer a useful approach to a positive optimal tax theory.

Introduction

> "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."
>
> Roy Blakey, *American Economic Review*, 1913

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

In 1935, U.S. President Franklin Delano Roosevelt sought to substantially increase the progressivity of taxes. 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...."

*Harvard Business School and NBER, mweinzierl@hbs.edu. Thanks to Rafael Di Tella, Louis Kaplow, Benjamin B. Lockwood, Greg Mankiw, David Moss, Andreas Peichl, and Glen Weyl for valuable discussions.

In 2011, U.S. President Barack Obama also sought to increase top marginal income tax rates. 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 not because we begrudge those who've done well — we rightly celebrate their success. Instead, it's 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 James Mirrlees (1971), stranger still is the use of these principles at all, as modern tax theory has largely moved away from them in favor of an approach emphasizing social welfare maximization.

Two hundred years ago, however, the underlying normative principle would have been recognized immediately as Adam Smith's (1776) 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."

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 ability to pay is therefore an appropriate basis for taxation. Simply put, Smith endorses *benefit-as-ability*-based taxation. Richard Musgrave (1959) labeled this logic the "classical" view of benefit-based taxation, a label I adopt in this paper.

This classical view of benefit-based taxation was highly influential in the late 18th and early 19th centuries but waned as John Stuart Mill's (1871) purely ability-based reasoning gained favor and the canonical contributions of Erik Lindahl (1919) shifted the focus of benefit-based tax research.[1] 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.[2] 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 in his theory.[3] The classical benefit-based logic exerted little influence on the welfarist objective assumed in modern Mirrleesian theory.

[1] As the remarkable surveys by Edwin Seligman (1908) and Richard Musgrave (1959) make clear, benefit based reasoning was a prominent, at times leading, approach among tax theorists through the 19th century. William Petty (1677), in particular, anticipated Smith's view, and Hobbes, Hume, and Rousseau among others subscribed to it in some form.

[2] Lindahl himself viewed his theory as conditional in this way, a view criticized as untenable by Samuelson (1955).

[3] 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 (2013) for a recent exploration of optimal taxation with preference heterogeneity.

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 (IFS, 2011) 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. Anthony Atkinson and Joseph Stiglitz's (1980) classic text and Louis 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 Bernard Salanié (2011), in his essential textbook on the economics of taxation, nor Robin Boadway (2012), in his excellent survey of optimal tax theory's implications for policy, mention Lindahl or benefit-based taxes.[4]

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 paper 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 paper 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 innate talent and public goods. Once public goods matter for ability, the classical *benefits-as-ability* view links seamlessly with the modern model. Thus, both first-best and constrained optimal benefit-based policy can be analyzed within the formal structure of modern tax theory and characterized using familiar methods.[5]

In particular, first-best policy according to this view of benefit-based taxation can be characterized in terms of simple and potentially observable elasticity parameters if we apply Lindahl's well-known method of measuring benefit. In this paper, 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 innate talent. 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 we find that optimal benefit-as-ability-based taxation resembles Mill's preferred "equal sacrifice" taxation, a possibility hinted at informally nearly forty years ago by Martin 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 characterize how different assumptions about the interaction of public goods and innate talent affect the progressivity of constrained optimal policy. To address the limitation that the classical benefit-based view does not provide a ranking of allocations other than the first-best, we specify a simple loss function that respects Pareto efficiency but allows the planner to choose the allocation, from the set of incentive compatible allocations, that deviates least from the first-best benefit-based allocation.

Why go to the trouble of attempting such a resuscitation?

The statements provided at the start of this paper suggest that the classical view of benefit-

[4] 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.

[5] 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 Henry Simons (1938) put it (quoted in Daniel Shaviro 2013).

based taxation is included in the criteria used to judge, or at least justify, tax policy in the United States. 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. 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.[6] 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.[7]

The second contribution of this paper 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 loss function used in the preceding first-best analysis. That loss 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 loss 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.[8] 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.

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

First, this paper 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 paper 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 paper does not claim to resolve the debate over whether the right approach to benefit-based taxation is Lindahl's. The proper way to capture the normative intuition behind benefit-based taxation has long been a topic of study, with prominent examples being the work of Aaron and McGuire (1970, 1976), Brennan (1976, 1979), West and Staaf (1979), Moulin (1987),

---

[6] See Weinzierl (2013a) and Saez and Stantcheva (2013); also see Zelenak (2006) for a different, legal theory, perspective. 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.

[7] See the discussion of normative diversity in Weinzierl (2013a). Not mentioned there, however, are two potentially interesting examples.

Richard Musgrave (1999) wrote: "Moreover, observers such as myself who tend to be egalitarian should not rule out the norm oof Lockean entitlement to earnings (Locke [1690]; 1960; Nozick 1974) as an alternative criterian 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."

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 necessaries 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."

[8] By narrowing its scope for accommodating multiple criteria, this approach substantially simplifies a more general technique introduced in Weinzierl (2013a) that combines normative criteria into a social loss function.

and Hines (2000).[9] Instead, I view Lindahl's approach as a natural way to formalize 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 most well-known example of this metaphor is usually attributed to U.S. Supreme Court Justice Oliver Wendell Holmes, who served on the Court from 1902 through 1932: "I like to pay taxes. With them I buy civilization."[10] In fact, Lindahl 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."

Third, the mixed normative objective utilized in this paper 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 Iván 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 modeled in alternative ways that are closer–in formal terms–to the standard approach (i.e., following the technique of Emmanuel Saez and Stefanie Stantcheva, 2013). This paper's approach chooses among Pareto-efficient allocations by assigning weights to a small number of normative principles with apparent empirical relevance.

The paper proceeds as follows. Section 1 formalizes the classical benefit-based view and analytically characterizes 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 innate talent. Section 2 quantitatively characterizes 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 gives 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.

[9] 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), criticized by Brennan (1976), and discussed by West and Straaf (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 paper thereby aligns with the arguments of Samuelson (1954, foonote 9). An alternative approach would be to tie taxes to the total, inframarginal benefits from public goods, perhaps along the lines of Moulin (1987).

[10] Holmes wrote a similar phrase into his decision on a 1927 case before the Court. It is worth recalling that, at the time, the U.S. 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.

1 Analytical characterization of classical benefit-based taxation

In this section I show how a single modification of the standard Mirrleesian optimal tax model enables us to formalize the classical view of benefit-based taxation and characterize optimal policy in terms of simple, potentially observable statistics.

1.1 The model

A population of individuals differ in their abilities to earn income w_i. In conventional models, these abilities are interpreted as innate and fixed relative to public goods.[11] The innovation in this paper is that individuals' income-earning abilities are functions of both innate talent and public goods funded by tax revenue.[12] Note that this paper takes only the first, most basic, step toward modeling 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.

Formally, $w_i = f(a_i, G)$, where $i \in I$ indexes innate talent types a_i, $G \geq 0$ is the level of spending on public goods, and $f(\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ 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

$$U(c_i, l_i) = u(c_i) - v\left(\frac{y_i}{f(a_i, G)}\right), \tag{1}$$

where c_i is private consumption for individual i and y_i is i's income, so that $\frac{y_i}{w_i}$ is work effort.

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

$$U_i^* = u(c_i^*) - v\left(\frac{y_i^*}{f(a_i, G^*)}\right) \tag{2}$$

for all $i \in I$.

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. I explore those questions in a companion paper, Weinzierl (2013b).

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 personalized tax "prices." The planner specifies this set of prices $\{\tau_i\}$, and individual i therefore faces the budget constraint

$$y_i - c_i - \tau_i pG \geq 0, \tag{3}$$

where p is the per-unit price 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.

[11] Much recent work in the Mirrleesian literature has considered other sources of endogeneity, such as human capital investment. See Kapicka (2006), Gelber and Weinzierl (2013), Best and Kleven (2013), Stantcheva (2014), among others.

[12] A fascinating 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.

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 personalized 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 section, I will follow Lindahl and focus on the first-best benefit-based policy. This is not merely for analytical tractability, though that is a side benefit. Rather, the first-best policy is the most direct guide to a normative principle's effects on policy, as evidenced by the analytical results below and how they compare to the first-best policy under a conventional utilitarian criterion (which is fully egalitarian). 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 will analyze numerically in the next section.

Lindahl's approach then has us consider a hypothetical scenario in which each individual i is allowed to choose her own consumption, work effort, and, importantly, level of public goods provision that maximize her utility subject to her 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.

$$u'(c_i) = \lambda \tag{4}$$

$$\frac{1}{f(a_i,G_i)} v'\left(\frac{y_i}{f(a_i,G_i)}\right) = \lambda \tag{5}$$

$$\frac{y_i f_G(a_i,G_i)}{f(a_i,G_i)} \frac{1}{f(a_i,G_i)} v'\left(\frac{y_i}{f(a_i,G_i)}\right) = \tau_i p \lambda \tag{6}$$

Lindahl defined optimal policy as that in which two conditions are satisfied: first, the personalized prices cause each type to prefer the same quantity of public goods; 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* $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ *is a First-Best Lindahl Equilibrium if and only if individuals maximize utility and the following conditions hold:*

$$G_i = G_{BB}^{FB} \ \forall i,$$

$$\sum_{i\in I} \tau_{i,BB}^{FB} = 1.$$

Next, I turn to characterizing optimal policy with conditions on its components, G_{BB}^{FB} and $\left\{\tau_{i,BB}^{FB}\right\}_i$ that depend on only (at least potentially) estimable statistics. All proofs of the following results are collected in the printed appendix at the end of this paper.

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

First, we 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, we can state the following result.

Proposition 1 *Samuelson Rule: If the policy* $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ *is a First-Best Lindahl Equilibrium, then the following condition is satisfied:*

$$\sum_{i \in I} \varepsilon_i^G\left(G_{BB}^{FB}\right) y_i = pG_{BB}^{FB}. \tag{7}$$

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).

The following corollary shows that this Samuelson rule 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:*

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

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

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

With these results, we have the following corollary of Proposition 1.

Corollary 1 *If the policy* $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ *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_{BB}^{FB} *satisfies:*

$$\varepsilon^G\left(G_{BB}^{FB}\right) \sum_{i \in I} y_i = pG_{BB}^{FB}. \tag{8}$$

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

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

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, $\varepsilon^G\left(G_{BB}^{FB}\right)$, toward funding the public goods that magnify innate talents. 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) labor 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 characterizing 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 innate talent, $\theta_i^{G,a}$, as:*

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

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 2014, and the citations therein). In this paper's context, it captures the degree to which public goods and innate talent magnify (or offset) each other in determining income-earning ability. If $\theta_i^{G,a} \leq 0$, innate talent 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 innate talent is positive but decreasing in the level of public goods; if $\theta_i^{G,a} > 1$, the elasticity of income-earning ability with respect to innate talent is increasing in the level of public goods. We can then state the following result.

Proposition 2 *Assume public goods and innate talent have positive partial effects on income-earning ability: i.e., $f_G(a_i, G) > 0$ and $f_a(a_i, G) > 0$. If $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ is a First-Best Lindahl Equilibrium, with corresponding incomes $\left\{y_{i,BB}^{FB}\right\}_i$ then the optimal average tax rate on individual i increases (decreases) with innate talent iff the Hicksian partial elasticity of complementarity between public goods and innate talent is greater than (less than) one. Formally, define the optimal average tax rate for individual i as $ATR_{i,BB}^{FB} = \tau_{i,BB}^{FB} p G_{BB}^{FB} / y_{i,BB}^{FB}$. Then,*

$$\frac{\partial ATR_{i,BB}^{FB}}{\partial a_i} > (<)\ 0 \;\Leftrightarrow \theta_i^{G,a}\left(G_{BB}^{FB}\right) > (<)\ 1.$$

The intuition for Proposition 2 is as follows. Public goods are assumed to magnify all abilities.[13] If the effect of public goods on ability has a greater elasticity to innate talent than does ability itself, that means that there is a complementarity between public goods and ability.[14] Thus, public

[13] In principle, this need not be the case, though if we 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.

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

goods spending is a source of greater benefits for the naturally talented, and optimal benefit-based taxation adjusts their taxes up accordingly.

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

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 $\mathbb{R}^+ \to \mathbb{R}^+$, 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 innate talent is the same no matter the level of public goods. With these results and Lemma 1, we have the following corollary of Proposition 2.

Corollary 2 *If $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ 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 $\mathbb{R}^+ \to \mathbb{R}^+$, then the optimal tax policy is a uniform average tax rate for all $i \in I$:*

$$ATR_{i,BB}^{FB} = \varepsilon^G\left(G_{BB}^{FB}\right)\ \forall\ i. \tag{10}$$

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

$$ATR_{i,BB}^{FB} = \gamma\ \forall\ i. \tag{11}$$

Of course, 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

John Stuart Mill's (1871) principle of equal sacrifice was the theory of optimal taxation that displaced Smith's classical benefit-based logic 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 Martin 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."

Using the model above, and assuming some familiar functional forms, we can show a formal connection that is consistent with this intuition.

To link these principles requires defining "sacrifice." To do so, we consider a hypothetical state in which public goods are provided at their optimal level, G_{BB}^{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:

$$u'\left(c_i^F\right) = \frac{1}{f\left(a_i, G_{BB}^{FB}\right)} v'\left(\frac{y_i^F}{f\left(a_i, G_{BB}^{FB}\right)}\right)$$

where G_{BB}^{FB} satisfies Proposition 1 and $c_i = y_i$ (there are no taxes) and the F superscript denotes the "free" public goods scenario.

To calculate utility levels, we need to specify the utility functional form. We assume the following familiar specification (which Mill 1871 appears to endorse in his discussion of marginal utility above a minimum income).

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

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

$$c_i^F = y_i^F = f\left(a_i, G^{BB}\right),$$

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

$$U\left(c_i^F, l_i^F\right) = \ln\left(c_i^F\right) - \frac{1}{\sigma}\left(\frac{y_i^F}{f\left(a_i, G_{BB}^{FB}\right)}\right)^{\sigma}. \tag{13}$$

With these, we can show the following result.

Proposition 3 *If $\left\{\left\{\tau_{i,BB}^{FB}\right\}_i, G_{BB}^{FB}\right\}$ 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 function differentiable function $h(a_i); \mathbb{R}^+ \to \mathbb{R}^+$ and $\gamma > 0$, then the First-Best Lindahl Equilibrium's utility levels are less than those in* (13) *by the same quantity for all individuals: that is, the policy generates "equal sacrifice" for all individuals.*

While this exact equivalence results holds only under a special set of functional form assumptions, the connection in Proposition 3 may have broader implications. In particular, it suggests that two of the most prominent alternatives to the conventional social welfare maximization criterion–the classical benefit-based logic of Smith (1776) and the equal sacrifice logic of Mill (1871)–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 favor of equal sacrifice taxation, nevertheless endorsed the same tax schedule as Smith and invoked Smith as inspiration for his reasoning.[15] 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 characterization of constrained–optimal (second-best) classical benefit-based taxes

A persistent critique of Lindahlian benefit-based reasoning (e.g., Mill 1871 and Samuelson 1954), is that individuals' benefits are unobservable. How to pursue benefit-based policy when constrained by this unobservability has been a topic of intense study for decades (e.g., Foley 1970). When one uses the classical view of benefit-based taxation, however, the problem of unobservable benefits is converted to the problem of unobservable ability–exactly the constraint at the heart of modern Mirrleesian tax theory. In this section, we take advantage of that feature of the classical view 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

[15] Both endorsed proportional taxation above a minimum level of income.

quantitatively to the way in which public goods and innate talent interact to produce income-earning ability. We 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 (2012); 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 a loss function that penalizes symmetric deviations from that principle's first-best (feasible but not necessarily incentive compatible) allocation.[16] I specify a very simple version of such a loss 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 loss function has a kink at its most-preferred allocation, so that the loss due to a deviation below that allocation is greater than the gain due to a similar-sized deviation above it.

Formally, recall the optimal feasible allocation that we derived and characterized in the previous section $\left\{c_{i,BB}^{FB}, y_{i,BB}^{FB}, G_{BB}^{FB}\right\}_i$. The loss function, denoted $\mathcal{L}$, is defined as follows.[17]

$$\mathcal{L}\left(\{c_i, y_i, G\}_i, \{c_{i,BB}^{FB}, y_{i,BB}^{FB}, G_{BB}^{FB}\}_i\right) = \sum_{i=1}^{I} V\left(U\left(c_i, \frac{y_i}{f(a_i, G)}\right), U\left(c_{i,BB}^{FB}, \frac{y_{i,BB}^{FB}}{f\left(a_i, G_{BB}^{FB}\right)}\right)\right), \tag{14}$$

where

$$V\left(U_i, U_{i,BB}^{FB}\right) = \begin{cases} -\delta\left(U_i - U_{i,BB}^{FB}\right) & \text{if } U_{i,BB}^{FB} < U_i \\ \left(U_{i,BB}^{FB} - U_i\right) & \text{if } U_{i,BB}^{FB} \geq U_i \end{cases}, \tag{15}$$
$$\text{for scalar } \delta:\ 0 \leq \delta \leq 1.$$

The loss function in expressions (14) and (15) applies weight δ, where $0 \leq \delta \leq 1$, to deviations of individual utility above the allocation $\left\{c_{i,BB}^{FB}, y_{i,BB}^{FB}, G_{BB}^{FB}\right\}_i$. 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).

With this loss function, we 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:*

$$\min_{\{c_i, y_i, G\}_{i=1}^{I} \in \{\mathbb{F} \cap \mathbb{IC}\}} \mathcal{L}\left(\{c_i, y_i, G\}_i, \{c_{i,BB}^{FB}, y_{i,BB}^{FB}, G_{BB}^{FB}\}_i\right), \tag{16}$$

[16] 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.

[17] Readers familiar with Weinzierl (2013a) will notice a strong similarity between the treatment of benefits-based taxation in this paper and Equal Sacrifice taxation in that paper. Though Section 1 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 a loss function that punishes deviations from an optimal allocation but respects Pareto efficiency.

where $\mathcal{L}$ is defined in expressions (14) *and* (15),
$\mathbb{F}$ *denotes the set of feasible allocations for the economy:*

$$\mathbb{F} = \left\{ \{c_i, y_i, G\}_{i=1}^{I} : \sum_{i=1}^{I} (y_i - c_i) \geq G \right\}, \tag{17}$$

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

$$\mathbb{IC} = \left\{ \{c_i, y_i, G\}_{i=1}^{I} : U(c_i, y_i/f(a_i, G)) \geq U(c_j, y_j/f(a_i, G)) \text{ for all } i, j \in \{1, 2, ..., I\} \right\}. \tag{18}$$

2.2 Parameterization 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 $\sigma = 3$.

A more important–and novel–question 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 paper does not base on direct empirical evidence its assumptions for how public goods affect ability, that is, 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), it is not this paper's purpose. 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). The 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 paper is to put that debate into formal terms, clarifying the implicit conditions that resolve that uncertainty and making it, at least in principle, an empirical question.

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 innate talent to be related to income-earning ability in a range of ways; and it ought to be as analytically simple as possible.

To address these challenges, I map individuals' innate types to fixed percentile positions in an ability distribution that is endogenous to the level of G.[18] Specifically, I will assume that income-earning ability is lognormally distributed and that the parameters of that distribution depend on the level of G.[19] 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, we can write:

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

where Φ is the standard normal cdf, $F(a_i)$ 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.

[18] The equality of opportunity literature, for instance Roemer (1998), has used a similar fixed-positions technique.

[19] 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 so-called double-Pareto-lognormal (DPLN) distributions. Future research may fruitfully extend this analysis to that case.

Income-earning ability can then be derived using the inverse cumulative distribution function and its parameters, so that for type i,

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

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:

$$\begin{aligned} \mu(G) &= \bar{\mu} - \frac{\xi}{G} + \ln G^{\gamma} \\ \sigma(G) &= \bar{\sigma} + \beta_1 G + \beta_2 G^2, \end{aligned} \tag{19}$$

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

Though simple, the forms assumed in (19) have a few appealing features.

First, this specification implies that all income-earning abilities approach zero as the state disappears (i.e., $\lim_{G \to 0} \mu(G) = -\infty$ because of the term $-\xi/G$). 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 (see Murphy and Nagel, 2003, 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 (19) implies that, for $\beta_1 = \beta_2 = 0$, the ability production function approaches a familiar form as ξ becomes small. That is,

$$\lim_{\xi \to 0} f(a_i, G) = \eta(a_i)\, G^{\gamma}, \tag{20}$$

where $\eta(a_i) = \exp\left(\bar{\sigma}\Phi^{-1}(F(a_i)) + \bar{\mu}\right)$ is constant. Expression (20) 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 we 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) was that a classical benefit-based logic would lead to just such a proportional tax.[20]

To demonstrate the range of potential optimal policies under this classical benefit-based criterion, 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 $\bar{\mu}$, ξ, 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 U.S. wage distribution.[21] For

[20] 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 paper, 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.

[21] The simulation has a distribution of individuals (indexed by their percentile position in the innate talent distribution) choose labor effort to maximize utility, taking into account the ability production function, flat tax rate, and share of G in output.

status quo tax policy, I follow Lockwood and Weinzierl (2013) and use an approximation of the current U.S. tax system with a flat tax rate of 40 percent and a lumpsum grant equal to the excess of tax revenue over government spending, which U.S. national accounts show to be approximately 7.6 percent of GDP. Data on reported wage rates from the National Longitudinal Survey of Youth (also discussed below) are best-described by a lognormal distribution with mean 2.67 and standard deviation 0.63.

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 (20). 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 innate talent. In the third column, I return to the baseline values of $\bar{\mu}$ and γ but relax the baseline assumption that β_1 and β_2 are zero and, instead, consider a case in which increases in G are relatively more beneficial for workers with high innate talent. Formally, to the extent that increases in G raise $\sigma(G)$, those higher in the talent 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.

Table 1: Parameterizations 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	0.076
$\bar{\sigma}$	0.74	0.74	0.295
β_1	0.00	0.00	0.065
β_2	0.00	0.00	-0.0015
Empirical targets under Status Quo tax policy			
$\mu(G) = 2.67$	2.67	2.67	2.67
$\sigma(G) = 0.63$	0.63	0.63	0.63
$G/\sum_i y_i$	0.076	0.076	0.076

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 U.S. tax policy, they apply quite different constrained-optimal policies.

To obtain the constrained-optimal policies, we simulate the planner's problem in expressions (16), (17), and (18). In the next section, we discuss in detail the parameterization of the loss function in this planner's problem. For this section, we assume $\delta = 0$, corresponding to a strict version of the classical benefit-based approach.

In Figure 1, we plot average tax rates along the ability distribution in the second-best optimal benefit-based policy for the three scenarios.

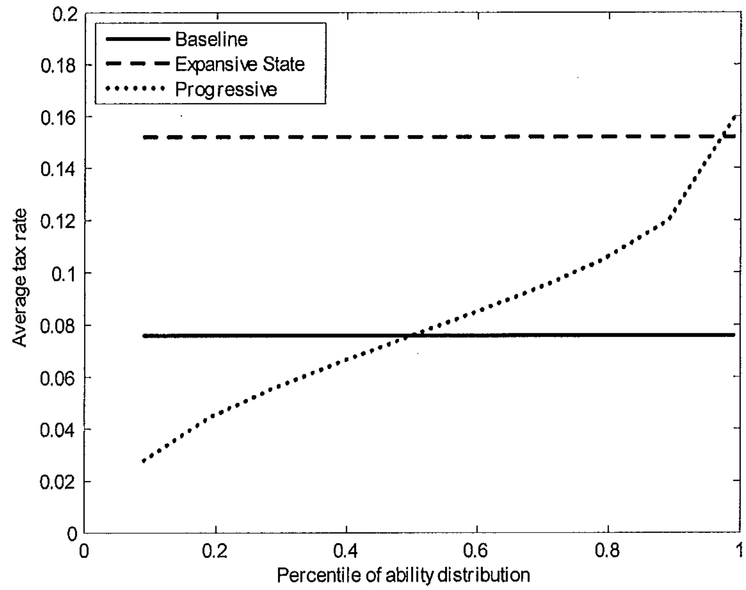

Figure 1: Average tax rates in the constrained-optimal policy under three sets of parameters from Table 1.

The shares of G in total income in these three optimal policies are 7.6 percent, 15.2 percent, and 11.8 percent. The message from these simulations is, therefore, 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. Benefit-based taxation is not necessarily "libertarian" in the sense of a minimal state; nor is it inconsistent with progressive taxation.[22] This flexibility is arguably behind the principle's long-standing and apparently widespread appeal.[23]

We can connect the results shown in Figure 1 to the analysis of first-best policy from Section 2 by calculating the realized values of the Hicksian partial elasticity of complementarity between public goods and innate talent, $\theta_i^{G,a}$ as defined in expression (9). In Table 3 we show these values for the three cases plotted in Figure 1:

[22]Both of these features are, in fact, consistent with Smith's view of taxation. Smith (1776) 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." Also see the earlier footnote on Smith's endorsement of progressivity.

[23]In other words, only if the principle had such flexibility would it be possible for 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, to invoke the same underlying principle as justification.

Table 3: Hicksian partial elasticities of complementarity $\theta_i^{G,a}$

Percentile of ability distrib.	1 (Baseline)	2 (Expansive State)	3 (Progressive)
10	1.00	1.00	2.60
30	1.00	1.00	1.73
50	1.00	1.00	1.54
70	1.00	1.00	1.45
90	1.00	1.00	1.49

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 innate talent 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 far from the only principle that may guide 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 the results of the previous section, that some principle other than the classical benefit-based logic must exert 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.[24]

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 formalized.

3.1 Formalizing a mixed objective function

To combine the classical benefit-based and utilitarian criteria into one objective, I take advantage of a convenient feature of the formalization of the classical benefit-based planner's problem in expressions (16), (17), and (18).[25] When $\delta = 1$, the planner's objective is equivalent to the conventional

[24] 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 policymakers 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 the following 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."

[25] Recall that, in that problem, the planner's desire to stay near the first-best benefit-based allocations is constrained not only by feasibility and incentive compatibility but also by Pareto efficiency. As discussed earlier, the Pareto efficiency requirement is incorporated into the loss function with which the planner evaluates any proposed allocation, in that the loss function not only punishes deviations of an individual's utility below the level he or she would obtain

utilitarian objective to maximize the unweighted sum of individual utilities.[26] In the previous section, we 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.

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. It turns out that these intermediate cases can yield a wide range of optimal policy results, including those that are consistent with several prominent features of existing policy.

3.1.1 Levels of progressivity and redistribution

First, I consider optimal average tax rate schedules for a calibrated U.S. ability distribution. I use fifty innate talent types for the simulations, such that each type represents two percent of the distribution $F(a_i)$. I use the "baseline" case from Table 1 for the values of $\bar{\mu}$, ξ, 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 results for four values of δ :

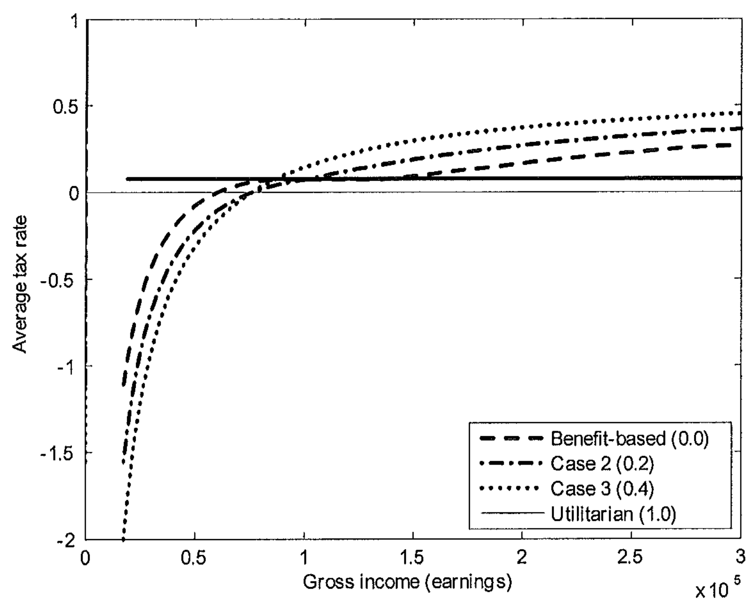

Figure 2: Average tax rates for four values of δ (in parentheses), assuming the baseline ability production function.

in the first-best benefit-based policy, but also *rewards* (at least weakly) deviations of individual utility above that level. Formally, in expression (15), the assumption that $\delta \geq 0$ means that Pareto efficiency is respected (at least in its weak form). Also recall that the loss function allows for asymmetric punishment and reward of utility deviations. Formally, the assumption that $\delta \leq 1$ means that symmetric deviations of utility from the benchmark allocation yield a net loss (or at least no net gain).

[26]To see this, impose $\delta = 1$ on expression (15) to obtain $V\left(U_{i,BB}^{FB}, U_i\right) = U_{i,BB}^{FB} - U_i$, so that the loss function in expression (14) simplifies to $\mathcal{L}(\cdot) = \sum\left(U_{i,BB}^{FB} - U_i\right)$, and the planner's objective (16) becomes $\min_{\{c_i,y_i,G\}_{i=1}^I \in \{\text{F}\cap\text{IC}\}} \sum\left(U_{i,BB}^{FB} - U_i\right)$. Because the allocations $U_{i,BB}^{FB}$ do not depend on the chosen allocations $\{c_i, y_i, G\}_{i=1}^I$, this objective is equivalent to: $\max_{\{c_i,y_i,G\}_{i=1}^I \in \{\text{F}\cap\text{IC}\}} \sum U_i$, the utilitarian planner's objective.

As this figure shows, the extent of redistribution in the optimal policy is substantially reduced when the classical benefit-based criterion receives greater weight in the objective function, under the baseline specification of the ability production function. Related, the average tax rates paid by high earners in these policies differ substantially, with rates of 8, 34, 41, and 48 percent.

In contrast to these large effects on redistribution and progressivity, this variation in the objective function has only negligible effects on the resulting distribution of income-earning abilities and level of public goods. This uniformity confirms that the classical benefit-based logic is consistent with a role for government in the economy similar to that implied by social welfare maximization.

3.1.2 Rank reversals

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 equalized across types, but those with higher income-earning ability are required to exert more labor effort. While the second-best optimal policy 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 (see, for example, King 1983 and Zelenak 2006, as well as the discussion in Saez and Stantcheva 2013).

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.

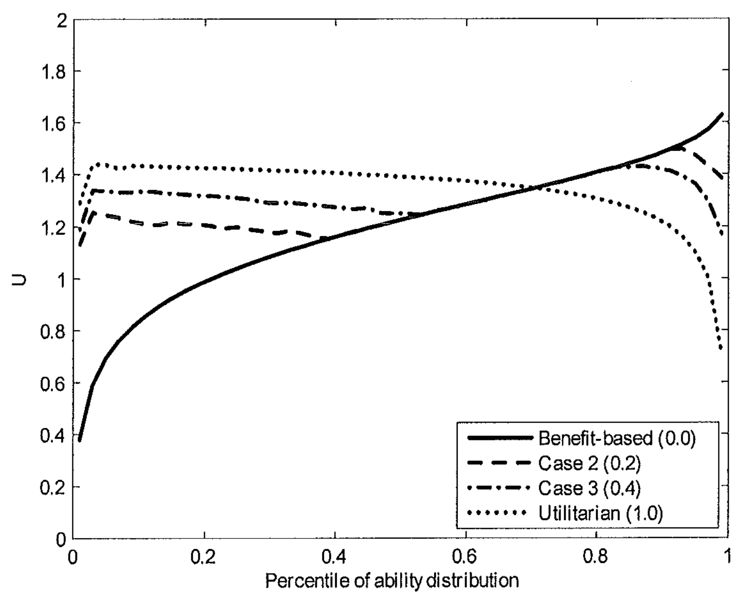

Figure 3: Utility levels by ability type in the first-best (full information) optimal allocations for four values of δ (shown in parentheses), assuming the baseline ability production function

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.

4 Conclusion

In this paper, I have explored whether we might incorporate a logic for tax design that appears to play a role in public reasoning but that has been largely set aside in modern tax theory. The idea that an individual ought to pay taxes based on the benefit he or she derives from the public goods the government provides has been, for many centuries, an intuitively compelling one. When that intuition is married to the view that the best measure of that benefit is a person's income-earning ability, benefit-based taxation turns out to fit seamlessly into the modern Mirrleesian approach. Applying Lindahl's approach to assigning taxes based on marginal benefits so defined, we found that first-best policy can be described analytically and second-best policy can be characterized using standard methods. In both cases, specifications of the model's specifics that are potentially estimable from standard data yield a variety of optimal policy results, including policies that match well several features of existing policy. To the extent that a mixture of this classical benefit-based reasoning and the more conventional welfarist (e.g., utilitarian) reasoning is a good approximation of prevailing objectives for tax policy, this model may offer a useful approach to positive optimal tax theory.

5 Appendix

5.1 Classical benefit-based reasoning in American political rhetoric

This brief section highlights additional instances in which prominent policymakers or policy commentators have relied on the classical logic of benefit-based taxation in their discussions of tax policy.

In the 2012 U.S. Presidential election, a fierce debate erupted over the following remarks made by President Obama.

> There are a lot of wealthy, successful Americans who agree with me — because they want to give something back. They know they didn't — look, if you've been successful, you didn't get there on your own....Somebody helped to create this unbelievable American system that we have that allowed you to thrive. Somebody invested in roads and bridges. If you've got a business — you didn't build that. Somebody else made that happen.

The Republican nominee for president, former Massachusetts Governor Mitt Romney, seized on the remark and responded:

> He [Obama] said this, "If you've got a business, you didn't build that. Somebody else made that happen." That somebody else is government, in his view. He goes on to describe the people who deserve the credit for building this business. And, of course, he describes people who we care very deeply about, who make a difference in our lives: our school teachers, firefighters, people who build roads. We need those things. We value school teachers, firefighters, people who build roads. You really couldn't have a business if you didn't have those things. But, you know, we pay for those things. Alright? The taxpayers pay for government. It's not like government just provides those to all of us and we say, "Oh, thank you government for doing those things." No, in fact, we pay for them and we benefit from them and we appreciate the work that they do and the sacrifices that are done by people who work in government. But they did not build this business.

The important feature of this exchange, for the purposes of this paper, is that Obama and Romney agree that public goods are essential for production and that the individuals who benefit from them ought to pay for them. Both, in other words, appear to support the classical view of benefit-based taxation.[27]

Examples of affinity for the classical view come not only from politicians and economists. Warren Buffett is a prominent American investor and policy advisor behind the so-called "Buffett Rule" that would set a floor on the average tax rate paid by high income earners. For many years, Buffett has argued for higher taxes on the rich with the metaphor he calls the Ovarian Lottery, as in the following statement.

[27] In fact, the disagreement between the candidates' positions appears still narrower when one considers that the Obama campaign later asserted that the President had been referring to roads and bridges, not the business, when he said "you didn't build that." That is, both Obama and Romney appear to grant individuals' some moral claim to their income. That position is consistent with the classical benefit based logic but in contrast with the conventional Mirrleesian model, as the latter is based on a Rawlsian framework in which individuals do not have such a claim–see Fleurbaey and Maniquet (2006) for more discussion.

"Imagine there are two identical twins in the womb, both equally bright and energetic. And the genie says to them, 'One of you is going to be born in the United States, and one of you is going to be born in Bangladesh. And if you wind up in Bangladesh, you will pay no taxes. What percentage of your income would you bid to be the one that is born in the United States?' It says something about the fact that society has something to do with your fate and not just your innate qualities. The people who say, 'I did it all myself,' and think of themselves as Horatio Alger – believe me, they'd bid more to be in the United States than in Bangladesh. That's the Ovarian Lottery."

Buffett's concept of the ovarian lottery reflects an affinity for both the Rawlsian normative view that lies behind the social welfare maximization approach of Mirrleesian optimal tax theory and the classical view of benefit-based taxation. First, the ovarian lottery is a close cousin of Rawls' veil of ignorance, and in other contexts Buffett directly endorses the idea that combining such a view with the assumption of diminishing marginal utility of income generates a reason–in fact *the* reason in the standard model–for income redistribution.[28] The remarkable aspect of Buffett's reasoning, for this paper's purposes, is that it also draws so directly on the benefit-as-ability logic for taxation. It seems that Buffett, like most people, has a multifaceted normative perspective on optimal tax design.

5.2 Proofs of Proposition 1 and Corollary 1

Use the second and third FOCs to set

$$\frac{y_i f_G(a_i, G_i)}{f(a_i, G_i)} \frac{1}{p} = \tau_i. \tag{21}$$

Use $\sum_{i \in I} \tau_i = 1$ to obtain

$$\sum_{i \in I} \frac{y_i f_G(a_i, G_i)}{f(a_i, G_i)} = p.$$

At the Lindahl optimum, $G_i = G_j = G^*$ for all types, so that

$$\sum_{i \in I} \frac{y_i f_G(a_i, G^*)}{f(a_i, G^*)} = p.$$

Now, note that the MRS for individual i when $G_i = G_j = G^*$ for all types is:

$$\frac{u'(c_i)}{\frac{y_i f_G(a_i, G^*)}{f(a_i, G^*)} \frac{1}{f(a_i, G^*)} v'\left(\frac{y_i}{f(a_i, G^*)}\right)} = \frac{1}{\tau_i p}$$

or, using the second FOC,

$$\frac{y_i f_G(a_i, G^*)}{f(a_i, G^*)} = \tau_i p.$$

Take the sum across i to obtain the result, imposing that $G_i = G_j = G^*$ for all types.

To prove the corollary, use the expression

$$f_G(a_i, G^*) = \gamma a_i^{\alpha} G^{\gamma - 1}$$

in the result from the Proposition and simplify.

[28]See, for example, Buffett's entry at the Giving Pledge: http://cms.givingpledge.org/. While Buffett, there, is speaking of voluntary redistribution, his advocacy for more progressive taxation suggests he would go beyond that.

5.3 Proofs of Proposition 2 and Corollary 2

Using the expression for individual taxes, we can express the average tax rate for individual i as:

$$ATR_i = \frac{\tau_i p G^*}{y_i} = \frac{f_G(a_i, G^*)}{f(a_i, G^*)} G^*. \tag{22}$$

Take the derivative with respect to a_i to obtain

$$\frac{\partial ATR_i}{\partial a_i} = \frac{f(a_i, G^*) f_{G,a}(a_i, G^*) - f_G(a_i, G^*) f_a(a_i, G^*)}{f(a_i, G^*) f(a_i, G^*) / G^*}. \tag{23}$$

The denominator is positive, so the sign is determined by the numerator. It is positive iff:

$$f(a_i, G^*) f_{G,a}(a_i, G^*) > f_G(a_i, G^*) f_a(a_i, G^*)$$

We assume all (single) partials are positive, so this can be written as

$$\frac{f_{G,a}(a_i, G^*)}{f_G(a_i, G^*)} > \frac{f_a(a_i, G^*)}{f(a_i, G^*)}$$

which gives the first condition. Now, the sign is negative iff

$$f(a_i, G^*) f_{G,a}(a_i, G^*) < f_G(a_i, G^*) f_a(a_i, G^*).$$

Again, assuming all single partials are positive, this yields the second condition:

$$\frac{f_{G,a}(a_i, G^*)}{f_G(a_i, G^*)} < \frac{f_a(a_i, G^*)}{f(a_i, G^*)}.$$

To prove the corollary, substitute the Cobb-Douglas function into the average tax rate expression and simplify.

5.4 Proof of Proposition 3

The optimal level of public goods is $G^* = \frac{\gamma}{p} \sum_{i \in I} y_i$. In the hypothetical allocation with no taxes, the individual sets

$$c_i = y_i = h(a_i)(G^*)^\gamma = h(a_i) \left(\frac{\gamma}{p} \sum_{i \in I} y_i \right)^\gamma,$$

so utility is

$$U(c_i, l_i) = \ln \left(h(a_i) \left(\frac{\gamma}{p} \sum_{i \in I} y_i \right)^\gamma \right) - \frac{1}{\sigma}. \tag{24}$$

In the actual optimal allocation, with costly public goods, Corollary 2 showed that $ATR_i = \tau_i p G^* / y_i = \gamma$. Individuals set

$$\begin{aligned}
\frac{1}{c_i} &= \lambda \\
\frac{1}{h(a_i) G^\gamma} \left(\frac{y_i}{h(a_i) G^\gamma} \right)^{\sigma - 1} &= \lambda \\
\gamma \frac{1}{G} \left(\frac{y_i}{h(a_i) G^\gamma} \right)^\sigma &= \tau_i p \lambda
\end{aligned}$$

Combining the second and third,

$$y_i = \frac{\tau_i p G^*}{\gamma}$$

Combining this with the first and second,

$$\frac{1}{c_i} = \left(\frac{1}{h\left(a_i\right)}\right)^{\sigma} \left(\frac{\tau_i p}{\gamma}\right)^{\sigma-1} G^{\sigma(1-\gamma)-1}$$

which, into the budget constraint, gives

$$G = \left(\frac{h\left(a_i\right)\gamma^{\frac{\sigma-1}{\sigma}}}{\tau_i p}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}\right)^{\frac{1}{1-\gamma}}.$$

Rearrange to isolate the tax share:

$$\tau_i = \frac{1}{p}\frac{h\left(a_i\right)\gamma^{\frac{\sigma-1}{\sigma}}}{G^{1-\gamma}}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}.$$

Then use the budget constraint on tax shares to solve for G

$$1 = \frac{1}{p}\frac{\gamma^{\frac{\sigma-1}{\sigma}}}{G^{1-\gamma}}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}\sum_{i\in I} h\left(a_i\right).$$

and thus

$$G = \left(\frac{1}{p}\gamma^{\frac{\sigma-1}{\sigma}}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{1}{1-\gamma}}.$$

so

$$\tau_i = \frac{h\left(a_i\right)}{\sum\limits_{i\in I} h\left(a_i\right)},$$

which is a nice side result that the tax share is the share of scaled innate talent. Now we can solve for allocations.

$$y_i = \frac{\tau_i p\left(\frac{1}{p}\gamma^{\frac{\sigma-1}{\sigma}}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}\sum\limits_{i\in I} h\left(a_i\right)\right)^{\frac{1}{1-\gamma}}}{\gamma}$$

$$c_i = \left(h\left(a_i\right)\right)^{\sigma}\left(\frac{\gamma}{\tau_i p}\right)^{\sigma-1}\left(\frac{1}{p}\gamma^{\frac{\sigma-1}{\sigma}}\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\sigma}}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{1-\sigma(1-\gamma)}{1-\gamma}}$$

and utility is

$$U\left(c_i, l_i\right)^* = \ln\left(h\left(a_i\right)\left(\frac{\gamma}{p}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{\gamma}{1-\gamma}}\left(\left(\frac{1}{1-\gamma}\right)^{\frac{1}{\sigma}}\right)^{\frac{1-\sigma(1-\gamma)}{1-\gamma}}\right) - \frac{1}{\sigma}\frac{1}{1-\gamma}. \tag{25}$$

$$U\left(c_i, l_i\right)^* = \ln\left(h\left(a_i\right)\left(\frac{\gamma}{p}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{\gamma}{1-\gamma}}\left(1-\gamma\right)^{1-\frac{1}{\sigma(1-\gamma)}}\right) - \frac{1}{\sigma}\frac{1}{1-\gamma}. \tag{26}$$

vs.

$$U\left(c_i, l_i\right) = \ln\left(h\left(a_i\right)\left(\frac{\gamma}{p}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{\gamma}{1-\gamma}}\left(\left(\frac{1}{1-\gamma}\right)^{\frac{1}{\sigma}}\right)^{\frac{\gamma}{1-\gamma}}\right) - \frac{1}{\sigma}. \tag{27}$$

Sacrifice is, therefore:

$$\begin{aligned} U\left(c_i, l_i\right) - U\left(c_i, l_i\right)^* &= \ln\left(h\left(a_i\right)\left(\frac{\gamma}{p}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{\gamma}{1-\gamma}}\left(\left(\frac{1}{1-\gamma}\right)^{\frac{1}{\sigma}}\right)^{\frac{\gamma}{1-\gamma}}\right) - \frac{1}{\sigma} \quad (28)\\ &\quad - \ln\left(h\left(a_i\right)\left(\frac{\gamma}{p}\sum_{i\in I} h\left(a_i\right)\right)^{\frac{\gamma}{1-\gamma}}\left(\left(\frac{1}{1-\gamma}\right)^{\frac{1}{\sigma}}\right)^{\frac{1-\sigma(1-\gamma)}{1-\gamma}}\right) - \frac{1}{\sigma}\frac{1}{1-\gamma} \quad (29) \end{aligned}$$

or

$$U\left(c_i, l_i\right) - U\left(c_i, l_i\right)^* = \frac{\sigma - 1}{\sigma}\ln\left(\frac{1}{1-\gamma}\right) + \frac{1}{\sigma}\frac{\gamma}{1-\gamma} \tag{30}$$

which is constant across types.

References

[1] Aaron, Henry and M. McGuire, (1970). "Public Goods and Income Distribution," *Econometrica*, 38.

[2] Aaron, Henry and M. McGuire (1976), "Reply to Geoffrey Brennan," *Econometrica*, 44.

[3] Atkinson and Stiglitz (1980) *Lectures in Public Economics*

[4] Best, Michael and Henrik Kleven (2013), "Optimal Income Taxation with Career Effects of Work Effort," Working Paper.

[5] Blakey, Roy G. (1914). "The New Income Tax," *American Economic Review*, 4(1), March.

[6] Boadway, Robin (2012). *From Optimal Tax Theory to Tax Policy*, MIT Press.

[7] Bovenberg, L. and B. Jacobs, (2011)."Optimal Taxation of Human Capital and the Earnings Function," *Journal of Public Economic Theory*, 13(6).

[8] Brennan, H. G. (1976). ""On the Distributional Implications of Public Goods," *Econometrica*, 44.

[9] Brennan, H. G. (1976). "A Rejoinder to the Aaron-McGuire Reply," *Econometrica*, 44.

[10] Buchanan, James and Richard Musgrave (1999). *Public Finance and Public Choice: Two Contrasting Visions of the State*, CESifo

[11] Feldstein, Martin (1976). "On the Theory of Tax Reform," *Journal of Public Economics*, 6, pp. 77-104.

[12] Fleurbaey and Maniquet (2006). "Fair Income Tax," *Review of Economic Studies*, 73.

[13] Foley, Duncan (1970). "Lindahl's Solution and the Core of an Economy with Public Goods," *Econometrica*, 38.

[14] Gelber, Alexander and Matthew Weinzierl (2014), "Equalizing Outcomes and Equalizing Opportunities: Optimal Taxation when Children's Abilities Depend on Parents' Resources," NBER Working Paper 18332.

[15] Institute of Fiscal Studies, Mirrlees, J.A., S. Adam, T. Besley, R. Blundell, S. Bon, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba (eds.) (2010) *Dimensions of Tax Design: The Mirrlees Review*. Oxford University Press.

[16] Kapicka, Marek. 2006. "Optimal Income Taxation with Human Capital Accumulation and Limited Recordkeeping." *Review of Economic Dynamics* 9(4): 612-639.

[17] Kaplow, Louis (2008). *The Theory of Taxation and Public Economics*, Princeton.

[18] Kaplow, Louis and Steven Shavell (2001). "Any non-welfarist Method of Policy Assessment Violates the Pareto Principle," *Journal of Political Economy* 109(2).

[19] Lindahl, Erik (1919). "Just Taxation – A Positive Solution" (E. Henderson, transl.) in Musgrave, Richard A. and A.T. Peacock, Eds, *Classics in the Theory of Public Finance*. Macmillan.

[20] Lockwood and Weinzierl (2013). "De Gustibus non est Taxandum," NBER Working Paper 17784.

[21] King, Mervyn (1983). "An Index of Inequality: With Applications to Horizontal Equity and Social Mobility," *Econometrica*, 51.

[22] Mankiw and Weinzierl (2009). "The Optimal Taxation of Height," *American Economic Journal: Economic Policy.*

[23] Matsumoto, Mutsumi, (2001), "Public Input Provision in an Optimal Income Tax Model," *FinanzArchiv.*

[24] Mill, John Stuart (1871). *Principles of Political Economy.* Oxford University Press, 1994. Book V, page 169.

[25] Mirrlees, J.A., (1971). "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies* 38, 175-208.

[26] Moulin, Herve (1987). "Egalitarian-Equivalent Cost Sharing of a Public Good," *Econometrica*, 55.

[27] Murphy, Liam and Thomas Nagel (2002), T*he Myth of Ownership,* Oxford.

[28] Musgrave, Richard A. (1959). *The Theory of Public Finance.* McGraw-Hill.

[29] Nozick, Robert (1974). *Anarchy, State, and Utopia.* New York: Basic Books.

[30] Petty, William (1677). "A Treatise of Taxes and Contributions," cited in Musgrave (1959).

[31] Roemer (1998). *Equality of Opportunity*, Harvard.

[32] Saez, Emmanuel and Stefanie Stantcheva (2013). "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." Working paper.

[33] Salanié, Bernard, (2011). *The Economics of Taxation*, MIT.

[34] Samuelson, Paul A, (1954). "The Pure Theory of Public Expenditure," *Review of Economics and Statistics.*

[35] Samuelson, Paul A, (1955). "Diagrammatic Exposition of a Theory of Public Expenditure," *Review of Economics and Statistics*, 37(4), November.

[36] Seligman, Edwin R. A. (1908), "Progressive Taxation in Theory and Practice," *American Economic Association Quarterly.*

[37] Shaviro, Daniel (2013). "The Forgotten Henry Simons," NYU Law School Working Paper.

[38] Simons, Henry (1938). *Personal Income Taxation*, University of Chicago.

[39] Smith, Adam (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.*

[40] Stantcheva, Stefanie, (2014), "Optimal Taxation and Human Capital Policies over the Life Cycle," Working paper.

[41] Stiglitz, Joseph (1987). "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," H*andbook of Public Economics*, Chapter 15.

[42] Weinzierl, Matthew (2013a). "The Promise of Positive Optimal Taxation," NBER Working Paper 18559, June.

[43] Weinzierl, Matthew (2013b). "Optimal Taxation when Government Matters." Working paper.

[44] Werning, Iván (2007). "Pareto Efficient Income Taxation," MIT Working paper.

[45] Zelenak, Lawrence (2006). "Taxing Endowment," *Duke Law Journal*, 55(6), April.