Restoring Ramsey tax lessons to Mirrleesian tax settings: Atkinson-Stiglitz and Ramsey reconciled*

Helmuth Cremer
Toulouse School of Economics
(University of Toulouse and Institut universitaire de France)
Toulouse, France

Firouz Gahvari
Department of Economics
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

May 2012, revised February 2016

*This paper has been presented in seminars at CESifo, Uppsala University, the University of Cologne and University Nova (Lisbon). We thank all the participants and in particular Felix Biebrauer, Sören Blomquist, Denis Epple, Kenneth Kletzer, Etienne Lehmann and Catarina Reis for their helpful suggestions and comments. We also thank Jean-Tirole for a number of inspiring discussions about the underlying problem.

Abstract

This paper restores many of the Ramsey tax/pricing lessons considered as outdated by the optimal tax approach to the Atkinson and Stiglitz (1976) framework wherein differential commodity taxes are considered to be redundant. The key to this finding is incorporating a "break-even" constraint for public firms into the Atkinson and Stiglitz framework. Break-even constraints are fundamental to the regulatory pricing literature but have somehow been overlooked in the optimal tax literature. This reconciles the optimal-tax and the regulatory-pricing views on Ramsey tax/pricing rules.

JEL classification: H2, H5.

Keywords: Ramsey taxes and pricing, break-even constraints, optimal tax literature, Atkinson and Stiglitz framework, regulatory pricing literature.

1 Introduction

In a recent contribution, Stiglitz (2015) discusses the importance of Frank Ramsey's 1927 classic paper on optimal taxation; referring to it as "brilliant" and a paper that "can be thought of as launching the field of optimal taxation and revolutionising public finance" (p. 1). Nevertheless, Stiglitz goes on to conclude his paper's Introduction by this observation "... later analyses showed crucial qualifications, so that the policy relevance of Ramsey's analysis may be limited". This paper calls for a reexamination of these "qualifications". They, we shall argue, stem from a Mirrleesian approach to optimal taxation that ignores public firms' break-even constraints. Yet, in practice, regulation is almost always associated with budget balancing requirement. A fact that has not escaped the attention of regulatory pricing literature forming a cornerstone of regulatory economics.[1]

Prior to Mirrlees (1971), the Ramsey tax framework served as a cornerstone of the optimal tax theory. The central question in this literature was that of designing (linear) commodity taxes to collect a given tax revenue. Labor income went generally untaxed or subjected to a linear tax. The main point made in this literature was that, except in very special cases, commodity taxes should not be uniform and that they should be set to balance efficiency and equity considerations. The efficiency aspects manifest themselves most clearly in the inverse-elasticity rules (derived when demand functions, Hicksian or Marshallian are separable). These efficiency-driven tax rules entail a regressive bias in that goods with low price-elasticities are often necessities consumed proportionately more by poorer households. This bias is then mitigated by the equity terms in the tax rules (often appearing as covariance terms). These tend to increase the tax rate on

[1] It is a mystery as to why it took the public finance literature more than four decades and until early 1970s to appreciate Ramsey's (1927) insights. It started with the seminal paper by Diamond and Mirrlees (1971) and followed by literally hundreds of papers. Baumol and Bradford (1970) and Sandmo (1976) provide an interesting history of this subject. In the original Ramsey problem, individuals are alike and there is no income tax. With heterogeneous individuals, one also allows for a uniform lump-sum tax or rebate (and possibly a linear tax on labor income); see Diamond (1975).

goods that are consumed proportionately more by richer households.

The Mirrleesian approach posed a serious challenge to the Ramsey tax framework approach and the lessons drawn from it. Mirrlees (1971) argued that the existence or absence of tax instruments must be rationalized on the basis of the informational constraints in the economy. This approach turns the Ramsey tax framework on its head by making nonlinear income taxation the most powerful tax instrument at the disposal of the government. In turn, the reliance on the nonlinear income tax has a devastating implication for the usefulness of commodity taxes. In their classic contribution, Atkinson and Stiglitz (1976) show that under some conditions—weak separability of preferences in labor supply and goods—an optimal nonlinear income tax is sufficient to implement any incentive compatible Pareto-efficient allocation. In other words, commodity taxes are redundant (or should be uniform). The Ramsey results come about, it is thus argued, merely as an artifact of restricting the income tax to be linear (an ad hoc and inconsistent assumption given the assumed informational structure). The Atkinson and Stiglitz (AS) result has had a tremendous effect in shaping the views of public economists concerning the design of optimal tax systems: In particular that prices should not be used for redistribution (even in a second best setting), and that in-kind transfers are not useful.[2]

A prominent public economics specialist summarizes his and the prevailing view in the literature this way: "Once we add the nonlinear income tax, if we keep homogeneous preferences and weak separability of labor, we have the following generalized view: The equity-efficiency tradeoff associated with income taxation is addressed with the

[2] It is by now well understood though that the AS result also has its own limitations. In particular, it may not hold under uncertainty (Cremer and Gahvari, 1995) or under multi-dimensional heterogeneity (Cremer, Gahvari and Ladoux, 1998, and Cremer, Pestieau and Rochet, 2001) and redistribution through prices may then once again be second-best optimal (Cremer and Gahvari, 1998 and 2002). Still, these limitations notwithstanding, it is fair to say that the Ramsey approach to taxation is considered as dated and no longer "state of the art". It does continue to occupy a prominent place in all advanced textbooks, but it is taught mainly as an introduction to tax design and not because of its practical relevance.

income tax (as done by Mirrlees), and we should follow Economics 101 first-best rules in other realms: uniform commodity taxes (so as not to distort consumption), public goods according to the Samuelson rule, marginal cost pricing, and so forth" (private correspondence).

This strong view notwithstanding, the Ramsey-type rules have had a more or less independent second—or some might argue first—life as a model of regulatory pricing.[3] In his pioneering paper that appeared prior to Mirrlees (1971), Boiteux (1956) studied linear pricing of a regulated multi-product monopoly that has to cover some "fixed cost" (for instance the infrastructure cost of the network). This is to be achieved through markups on the monopoly's different products (equivalent to taxes).[4] Formally, this problem is equivalent to a Ramsey tax model with the fixed cost playing the role of the government's tax revenue requirement.

To sum, optimal tax and regulatory pricing literatures appear to have diverged in the way they view the practical relevance of the AS result (see Section 2.1 below). Whereas Ramsey pricing is considered totally passé in the optimal tax literature (even by Econ 101 standard according to our earlier quote from a prominent expert), Ramsey-type lessons permeate the field of regulatory economics. This is quite surprising considering the fact that the issues the two literatures address have an identical formal structure. In both cases, there is a public authority whose objective is to raise a fixed amount of revenue (where the revenue finances the government's expenditures in the optimal tax literature and the firms' fixed costs in the regulation literature).[5] This paper is thus also an attempt to reconcile these divergent views and bring them together. To this

[3] While it is true that regulators and especially competition authorities are often reluctant to accept Ramsey pricing arguments, this is *not* because of the AS result. Their objection is more of legal and procedural nature. In particular, Ramsey prices are often viewed as "discriminatory" and subject to informational problems when the operator is better informed about demand condition than the regulator.

[4] We follow the terminology used in the regulation literature but, in reality, this is a quasi-fixed cost relevant also in the long-run (a non convexity in the production set).

[5] Redistributive concerns are not confined to the tax literature and also appear in the regulatory economics literature.

end, we incorporate the regulatory economics focus on budget balance into a Mirrleesian optimal tax framework. As far as we know, a comprehensive analysis of this sort has not been attempted before.

The optimal tax setting we consider combines nonlinear income taxation with linear taxation/pricing of consumption goods. The informational structure that underlies this setting is now standard in the Mirrleesian optimal taxation literature. First, individuals' earning abilities and labor supplies are not publicly observable, but their pre-tax incomes are. This rules out type-specific lump-sum taxes but allows for nonlinear income taxation. Second, individual consumption levels of goods, whether subject to regulation or not, are not observable so that nonlinear taxation and/or pricing of goods are not possible. On the other hand, anonymous transactions are observable making linear commodity taxation feasible.

To account for the regulatory economics balanced budget concerns, we assume that a subset of the goods are produced by a public or regulated firm that has to cover a fixed cost through markups on the different commodities it sells.[6] This latter constraint gives rise to a "break-even" constraint on the part of the firm. This comes on top of the overall government's budget constraint. In other words, while our setting is that of AS, we depart from the existing optimal tax literature by formally incorporating a binding break-even in our model. Break-even constraints are fundamental to the regulatory pricing literature; yet somehow they have been overlooked in the optimal tax literature. We draw on this literature to show that there exist good justifications for imposing such a constraint. Some are due to informational issues pertaining to the opportunity to invest in a given sector or the incentives to reduce costs.[7]

While we optimize over all tax instruments including income, we are not concerned

[6] Whether this firm is public, as in Boiteux's world, or private but regulated does not matter. Either way, one implicitly assumes that the firms' revenues must also cover some "fair" rate of return on capital.

[7] We will not formally model these type of informational asymmetries in this paper. However, we shall return to this issue in the Conclusion and provide some indication on how our results would be affected if this extra source of information asymmetry were considered.

with the properties of the income tax schedule. Our aim is to solely study the commodity taxation and pricing rules (for goods produced by public/regulated firms as well as those produced subject to no regulation). We derive these rules for general preferences but concentrate on the case of weakly-separable preferences between labor supply and goods that underlies the AS result. We shall refer to this environment as the AS setting/framework/model and contrast it with the Ramsey environment wherein all tax instruments are linear. The fundamental contribution of our paper is that the AS setting with a break-even constraint restores many of the traditional Ramsey tax/pricing features that have been questioned by modern optimal tax theory.

Specifically, we demonstrate that, given a break-even constraint, not only is it desirable to tax the goods produced by the public/regulated firm but also other goods. Intuitively, taxation of privately-produced goods are generally needed to offset the distortions created by the public/regulated firm's departure from marginal cost pricing. This result stands in sharp contrast to AS result on the redundancy of commodity taxes. We then illustrate and elaborate on our findings by studying a simple framework with one publicly-provided and two privately-provided goods.

To interpret our results, and for comparison purposes, we also recast and derive the pricing rules for a setting wherein nonlinear income taxation is ruled out, and taxes and markups are used to finance *both* a revenue requirement and a fixed cost (rather than only one of the two as is traditionally done in Ramsey models). While this yields predictable results and is not of much interest in itself, it is a useful reference point. It encapsulates the effects of imposing a break-even constraint in a Ramsey setting and thus serves as a benchmark for understanding and interpreting our results (derived for the AS setting with a break-even constraint). We present this benchmark as an Appendix to the paper. Comparing the results derived in our setting to those in this benchmark highlights and isolates the role that commodity taxes do play in the AS setting (contrary to the AS's own "no role" result).

The two special cases of independent Hicksian and independent Marshallian demand curves provide further insights into the nature of the tax/pricing rules in our model. In the separable Hicksian demand case, we find that private goods (not included in the break-even constraint) should go untaxed. On the other hand, public firms should follow pricing rules that are purely efficiency-driven and Ramsey type: Goods are taxed inversely to their compensated demand elasticity regardless of their distributional implications. Redistribution is taken care of by the income tax (allowing public firms' prices to be adjusted for revenue raising as in the Ramsey model with identical individuals). This is to be contrasted with today's prevailing view—based on the AS framework that ignores break-even constraints—that commodity taxes are redundant. It also differs from earlier Ramsey pricing views that commodity taxes should follow inverse elasticity rules adjusted for redistributive concerns.

Results become less predictable in the case where Marshallian demands are independent. Here, allowing for a break-even constraint in the AS framework, resurrects a role for commodity taxes that go beyond the goods produced subject to the break-even constraint. Instead, it spills over to the taxation of other goods as well. One continues to get inverse elasticity rules as in the Ramsey model; however, their structure differs from the traditional expressions in the Ramsey model. On the one hand, they are more complicated than the pure efficiency rules. On the other hand, there is no covariance (or similar) term that captures redistributive considerations. Instead, they contain "tax revenue terms" that measure the social value of the extra tax revenues generated from demand variations that follow the (compensating) adjustments in disposable income. These terms lead to predictions that are similar to those coming from the many-household Ramsey model albeit without redistributive concerns; namely, that goods with higher demand elasticities should be taxed more heavily.

Finally, we study what is arguably the most celebrated general result obtained in the Ramsey model; namely, the (un)equal proportional reduction in compensated demands property. We show that, in contrast to the single-household Ramsey model, the reductions differ across goods. This in and of itself is not particularly surprising given the presence of heterogeneous households. More interestingly, compared to the many-household Ramsey model with the break-even constraint, we find that the redistributive considerations are once again replaced by tax revenue terms.

2 The model

There are H types of individuals, indexed $j = 1, 2 \ldots, H$, who differ in their wages, w^j, but have identical preferences over goods and leisure.[8] All goods are produced at a constant marginal cost which we normalize to one. Some, $\underline{x} = (x_1, x_2, \ldots, x_n)$, are produced by the private sector; the rest, $\underline{y} = (y_1, y_2, \ldots, y_m)$ are produced by a public or regulated firm which incurs a fixed cost. The firm is constrained to break even by marking up its marginal costs.[9] Let $\underline{p} = (p_1, p_2, \ldots, p_n)$ denote the consumer price of $\underline{x}$ and $\underline{q} = (q_1, q_2, \ldots, q_n)$ the consumer price of $\underline{y}$. Finally, denote the commodity tax rates on $\underline{x}$ by $\underline{t} = (t_1, t_2, \ldots, t_n)$ and the public firms' commodity-tax-cum-mark-ups by $\underline{\tau} = (\tau_1, \tau_2, \ldots, \tau_m)$. We have $p_i = 1 + t_i$ $(i = 1, 2, \ldots, n)$ and $q_s = 1 + \tau_s$ $(s = 1, 2, \ldots, m)$.

Individual consumption levels are not publicly observable but anonymous transactions can be observed. Consequently, commodity taxes must be proportional and public sector prices are linear. For the remaining variables, the information structure is the one typically considered in mixed taxation models; see e.g., Christiansen (1984) and Cremer and Gahvari (1997). In particular, an individual's type, w^j, and labor input, L^j, are not publicly observable; his before-tax income, $I^j = w^j L^j$, on the other hand, is. Consequently, type-specific lump-sum taxation is ruled out but non-linear taxation of incomes is feasible.

[8] Our results will not change if a continuous distribution of types are considered.

[9] Alternatively one can think of a privately owned regulated firm whose prices are set to cover cost plus a fair rate of return on capital.

To characterize the (constrained) Pareto-efficient allocations we derive an optimal revelation mechanism. For our purpose, a mechanism consists of a set of type-specific before-tax incomes, I^js, aggregate expenditures on private sector and public sector goods, c^js, and two vectors of consumer prices (same for everyone) $\underline{p}$ and $\underline{q}$ (for $\underline{x}$ and $\underline{y}$). To proceed further, it is necessary to consider the optimization problem of an individual for a *given* mechanism $(\underline{p}, \underline{q}, c, I)$. Formally, given any vector $(\underline{p}, \underline{q}, c, I)$, an individual of type j maximizes utility $u = u\left(\underline{x}, \underline{y}, I/w^j\right)$ subject to the budget constraint $\sum_{i=1}^{n} p_i x_i + \sum_{s=1}^{m} q_s y_s = c$. The resulting conditional demand functions for $\underline{x}$ and $\underline{y}$ are denoted by $x_i = x_i\left(\underline{p}, \underline{q}, c, I/w^j\right)$ and $y_s = y_s\left(\underline{p}, \underline{q}, c, I/w^j\right)$.[10] Substituting in the utility function yields the conditional indirect utility function

$$v\left(\underline{p}, \underline{q}, c, I/w^j\right) \equiv u\left[\underline{x}\left(\underline{p}, \underline{q}, c, I/w^j\right), \underline{y}\left(\underline{p}, \underline{q}, c, I/w^j\right), I/w^j\right].$$

Thus, a j-type individual who is assigned c^j, I^j will have demand functions and an indirect utility function given by

$$\begin{aligned} x_i^j &= x_i(\underline{p}, \underline{q}, c^j, I^j/w^j), && (1)\\ y_s^j &= y_s\left(\underline{p}, \underline{q}, c^j, I^j/w^j\right), && (2)\\ v^j &= v\left(\underline{p}, \underline{q}, c^j, I^j/w^j\right). && (3) \end{aligned}$$

Similarly, the demand functions and the indirect utility function for a j-type who claims to be of type k, the so-called mimicker, is given by

$$\begin{aligned} x_i^{jk} &= x_i(\underline{p}, \underline{q}, c^k, I^k/w^j), && (4)\\ y_s^{jk} &= y_s\left(\underline{p}, \underline{q}, c^k, I^k/w^j\right), && (5)\\ v^{jk} &= v\left(\underline{p}, \underline{q}, c^k, I^k/w^j\right). && (6) \end{aligned}$$

[10] These demand functions are derived conditional on a given I. Hence the conditional qualifier.

2.1 The break-even constraint

The information structure posited above describes only the informational asymmetries between the tax administration and the taxpayers typically assumed in the optimal tax literature. This does not rationalize a break-even constraint which is the missing link between the optimal tax and regulatory economics literatures. In settings where tax policy is restricted only by informational considerations of this type, break-even constraints could be undone by simple lump-sum transfers from the government to the operators. Yet, in practice, regulation is almost always associated with budget balancing requirements. Which explains why they form a cornerstone of regulatory economics.

2.1.1 Break-even constraints in regulatory economics

Ramsey-Boiteux (RB) pricing continues to play an important role in the sectors still subject to some form of regulation even though, over the last few decades, the scope of regulation has declined. A prominent example is the postal sector in the US where Ramsey-Boiteux pricing remains an important benchmark in regulatory hearings; see Crew and Kleindorfer (2011, 2012). As a matter of fact, not only has RB pricing kept its position but it has even found new applications in setting of regulatory reform and market liberalization. For instance, while the original Boiteux model concerns a monopoly, Ware and Winter (1985) show that generalized RB rules prevail in imperfectly competitive markets. Furthermore, Laffont and Tirole (1990) argue that one should price the network access an incumbent operator has to provide to its upstream competitors on the basis of RB logic.

Another interesting result, shown by Vogelsang and Finsinger (1979), is that Ramsey prices can be decentralized through an iterative procedure based on a global price cap; see also Laffont and Tirole (2000, page 67). More generally, in the literature on incentive regulation, Ramsey prices are viewed as a kind of "ideal" solution: they represent a so-called "full information" benchmark. One should bear in mind, however, that the informational asymmetries the incentive literature focuses on are between the regulator (public authority) and the firms. This is in contrast with the informational asymmetries between the tax administration (public authority) and the workers (taxpayers) that are at the heart of the optimal tax literature. Finally, the idea that prices ought to be used for redistributive purposes is the rational for a great deal of regulatory policies including social tariffs and more generally universal service requirements; see Cremer *et al.* (2001) for a discussion of the theoretical foundations of these policies and their practical implementation.

Now regulatory economists typically take the break-even requirement as an exogenously given constrained; it is neither questioned nor justified in the relevant literature. For instance Brown and Sibley (1986, p. 35) refer to the United States "*tradition of regulation: utilities' are expected to cover their costs*". Viscusi *et al.* (1998), who present an otherwise very comprehensive overview of regulatory and anti-trust policy, devote just a few lines to the break-even issue and take a somewhat cynical view by stating (p. 371): "*Regulators do not see as their primary objective achieving economic efficiency. Rather, they appear to seek a set of prices that are not unduly discriminatory but that permit total revenue to cover total cost.*"

2.1.2 Rationalizing break-even constraints

Going beyond this near axiomatic view of break-even constraints, there are both informational and non-informational grounds that can rationalize them. Political economy considerations provide an obvious example for the latter. Public authorities are typically reluctant to finance a structural deficit in a given sector through subsidies from the general budget. For example, voters may consider it to be more "fair" to have natural gas users pay for the transportation costs associated with the network of pipelines and pumping stations (as compared to the taxpayers). Legal issues provide another reason. In the EU, for instance, "State Aid" is illegal: Member states are not allowed to finance their operators' deficits through subsidies.[11]

The information-based rationalizations are not explicitly addressed in our setting. These are summarized by Laffont and Tirole (1993, pp 23–30). They give two basic arguments. One, which they ascribe to Coase (1945, 1946), applies to "a firm or a product whose existence is not a forgone conclusion". Unless an activity at least breaks even, one cannot be sure that its production is beneficial for the society to warrant the government covering its fixed costs. Roughly speaking, Coase argues that, absent a budget constraint, the government is entrusted to make this decision without having the appropriate information. With a break-even constraint, this decision is effectively made by consumers who thereby reveal if their willingness to pay for the product is sufficiently high. The second argument is related to incentives associated with informational asymmetries between the regulator and the firm (which we do not formally model).[12] For instance, Allais (1947) has argued that the absence of a break-even constraint would create "inappropriate incentives for cost reduction".

2.2 Constrained Pareto-efficient allocations

Denote the government's external revenue requirement by $\bar{R}$ and the fixed costs of public firms by F. Constrained Pareto-efficient allocations are described, indirectly, as follows.[13] Maximize

$$\sum_{j=1}^{H} \eta^j v\left(\underline{p}, \underline{q}, c^j, I^j/w^j\right) \tag{7}$$

[11] Unless the aid comes as a compensation for specific public-service obligations imposed on an operator. The State Aid legislation is in turn motivated by anti-trust considerations and specifically the concern that member states might subsidize their "national champion".

[12] We discuss this issue further in the Conclusion.

[13] Indirectly because the optimization is over a mix of quantities *and* prices. Then, given the commodity prices, utility maximizing individuals would choose the quantities themselves.

with respect to $\underline{p}, \underline{q}, c^j$ and I^j where η^js are constants with the normalization $\sum_{j=1}^H \eta^j = 1$.[14] The maximization is subject to the resource constraint

$$\sum_{j=1}^{H} \pi^j \left[(I^j - c^j) + \sum_{i=1}^{n} (p_i - 1) x_i^j + \sum_{s=1}^{m} (q_s - 1) y_s^j \right] \geq \bar{R} + F, \tag{8}$$

the break-even constraint

$$\sum_{j=1}^{H} \pi^j \left[\sum_{s=1}^{m} (q_s - 1) y_s^j \right] \geq F, \tag{9}$$

and the self-selection constraints

$$v^j \geq v^{jk}, \quad j, k = 1, 2, \ldots, H. \tag{10}$$

Denote the Lagrangian expression by $\mathcal{L}$, and the Lagrangian multipliers associated with the resource constraint (8) by μ, the public firms' break-even constraint (9) by δ, and with the self-selection constraints (10) by λ^{jk}. We have

$$\begin{aligned} \mathcal{L} &= \sum_j \eta^j v^j + \mu \left\{ \sum_j \pi^j \left[(I^j - c^j) + \sum_{i=1}^{n} (p_i - 1) x_i^j + \sum_{s=1}^{m} (q_s - 1) y_s^j \right] - \bar{R} - F \right\} \\ &\quad + \delta \left\{ \sum_j \pi^j \left[\sum_{s=1}^{m} (q_s - 1) y_s^j \right] - F \right\} + \sum_j \sum_{k \neq j} \lambda^{jk} (v^j - v^{jk}). \end{aligned} \tag{11}$$

The first-order conditions of this problem with respect to I^j, c^j, for $j, k = 1, 2, \ldots, H$, and p_i, q_s, for $i = 2, 3, \ldots, n$, and $s = 1, 2, \ldots, m$, characterize the Pareto-efficient allocations constrained both by the public firms' break-even constraint, the resource constraint, the self-selection constraints, as well as the linearity of commodity tax rates (see the paragraph below as to why the optimization does not extend to $i = 1$). They are derived in Appendix A.[15]

[14] The maximization must leave one of the prices out; see the discussion at the end of this Section.

[15] We assume that the second-order conditions are satisfied. Their violation is only interesting in conjunction with the properties of the income tax schedule and the question of bunching.

The reason that we do not optimize over p_1 is well-known in the optimal tax literature. With $\underline{x}$ and $\underline{y}$ being homogeneous of degree zero in $\underline{p}, \underline{q}$, and c, consumer prices can determined only up to a proportionality factor. Consequently, one of the consumer prices must be normalized. We choose p_1 and set its value at $p_1 = 1$.[16] Having stated this, a caveat is in order. In the absence of the break-even constraint, the normalization of one of the consumer prices is without any loss of generality. In our setting, the fact that consumer prices can be determined only up to a proportionality factor implies, at first glance, that the break-even constraint may be rendered inconsequential. This is the idea that the government can always raise all prices proportionately to cover the break-even constraint of the public firm! However, imposing a binding break-even constraint also rules out the possibility of an across the board uniform increase in *all* consumer prices, including those of the public firm, for the revenues of the public firm to cover its fixed costs. Indeed, such an across the board increases in all prices is in contradiction with the very idea of imposing a break-even constraint in the first place. This "policy" will never work when there are multiple firms with different cost structures.[17]

3 Atkinson and Stiglitz theorem and optimal commodity taxes

In the standard mixed taxation model without the break-even constraint, assuming preferences are weakly separable in goods and labor supply, the Atkinson and Stiglitz (1976) theorem on the redundancy of commodity taxes holds. The particular feature of separability that drives the AS result is the property that a j-type who pretends to be

[16] Observe that with $p_1 = 1$, one can replace $\sum_{i=1}^{n}(p_i - 1)x_i^j$ in (11) with $\sum_{i=2}^{n}(p_i - 1)x_i^j$.

[17] Suppose there are two public firms. Refer to the one which has a relatively higher fixed cost as #1 and the other as #2. Raising prices proportionately to balance the budget of #1 must imply that #2 will have a surplus, while raising them to balance the budget of #2 must imply that #1 will have a deficit.

of type k will have the same demand as type k. That is,

$$x_i^{jk} = x_i^k = x_i(\underline{p}, \underline{q}, c^k), \tag{12}$$

$$y_s^{jk} = y_s^k = y_s\left(\underline{p}, \underline{q}, c^k\right). \tag{13}$$

This arises because with weak-separability, preferences take the following form $u = u\left(f\left(\underline{x}, \underline{y}\right), I/w^j\right)$. Under this circumstance, the (conditional) demand functions for $\underline{x}$ and $\underline{y}$ specified in equations (1)–(2) and (4)–(5) will be independent of I/w^j so that $x_i = x_i\left(\underline{p}, \underline{q}, c\right)$ and $y_s = y_s\left(\underline{p}, \underline{q}, c\right)$. Moreover, the indirect utility function too will be weakly separable in $\left(\underline{p}, \underline{q}, c\right)$ and I/w^j and written as $v\left(\phi\left(\underline{p}, \underline{q}, c\right), I/w^j\right)$.

The above property has far reaching implications for optimal commodity taxes in our setting too; both those produced by the public firm as well as privately. To derive these, introduce the compensated version of demand functions (1)–(2). Specifically, denote the compensated demand for a good by a "tilde" over the corresponding variable. Let Δ denote the $(n+m-1) \times (n+m-1)$ matrix derived from the Slutsky matrix, aggregated over all individuals, by deleting its first row and column,

$$\Delta = \begin{pmatrix} \sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial p_2} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial q_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial p_2} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial q_m} \\ \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial p_2} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial p_2} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial q_m} \end{pmatrix}. \tag{14}$$

We prove in Appendix A that in this case optimal commodity taxes satisfy the following equations,[18]

$$\begin{pmatrix} t_2 \\ \vdots \\ t_n \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \\ \vdots \\ \left(1+\frac{\delta}{\mu}\right)\tau_m \end{pmatrix} = -\frac{\delta}{\mu}\Delta^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sum_j \pi^j y_1^j \\ \vdots \\ \sum_j \pi^j y_m^j \end{pmatrix}. \tag{15}$$

The implication of equation (15) stands in sharp contrast to the prevailing view in the literature as summarized in the quote we gave in the Introduction. It demonstrates quite clearly that, notwithstanding the Atkinson and Stiglitz (1976) theorem, $\underline{t} = \underline{0}$ and $\underline{\tau} = \underline{0}$ are not a solution to (15). Commodity taxes and departures from marginal cost pricing are necessary components of Pareto-efficient tax structures. The so-called "Economics 101" intuition appears to be a false one!

Equally crucial is to realize that the underlying reason for this result is the existence of the "break-even" constraint. To see this, observe that if $F = 0$, i.e. if there is no fixed cost, the break-even constraint becomes irrelevant and thus non-binding in our optimization problem. Under this circumstance, $\delta = 0$ and the right-hand side of (15) is reduced to a vector of $n+m-1$ zeros. One then obtains $t_i = 0$ and $\tau_s = 0$ (marginal cost pricing) for all $i = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, m$, and returns to the Atkinson and Stiglitz result that commodity taxes are redundant. With $F > 0$, the break-even constraint is necessarily violated under marginal cost pricing so that $\delta > 0$. In this case, the first n elements of the vector in the right-hand side of (15) continue to be zero, but the other m elements differ from zero. It then follows that the solution no longer implies all ts and all τs are zero. Nor will it be the case that the ts are necessarily zero. This point, that the existence of a break-even point requires not only the taxation of goods produced by the public firm but also the taxation of privately-provided goods, constitutes the major lesson of our study.

[18] Observe that Δ is of full rank so that its inverse exists; see Takayama (1985).

The second important lesson we are seeking to answer is the extent to which the pricing rules defined by (15) resemble the traditional Ramsey rules. Traditionally, however, the Ramsey pricing rules are derived for either a unified government budget constraint (in the public finance literature), or for a public firm (in the regulation literature), but not for the two together as we have done here. Adding on a break-even constraint to the government's budget constraint in the Ramsey problem, however, does not change the structure of Ramsey taxes/pricing rules. This is easy to show. For completeness, and to establish a benchmark for comparison, we do this and report it in Appendix B.

The counterpart to (15) in the Ramsey model with break-even constraint is equation (B8) in Appendix B. The most striking difference between the two sets of results is the lack of any distributional considerations (15).[19] This constitutes the second important lesson of our study: The tax/pricing rules for both types of goods, those that are produced privately and those that are provided through the public firm, are not affected by redistribution concerns. It is important to note that this statement concerns tax *rules* as opposed to tax *levels* which will obviously be affected.

To gain further insights into the nature of commodity taxes in (15), we next resort to a simple special case with one private-sector and one public-sector good.

3.1 Two privately-produced goods, one public

Under this simple structure, and with $t_1 = 0$, t_2 and τ_1 are found from equation (15) to be

$$\begin{pmatrix} t_2 \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \end{pmatrix} = \frac{-\delta}{\mu\left[\sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial p_2} \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_1} - \left(\sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial q_1}\right)^2\right]} \begin{pmatrix} -\sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial q_1} \sum_j \pi^j y_1^j \\ \sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial p_2} \sum_j \pi^j y_1^j \end{pmatrix}.$$

[19] They appear in (B8) through γ^j.

It immediately follows from the above that

$$t_2 = \frac{\sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial q_1}}{-\sum_j \pi^j \frac{\partial \tilde{x}_2^j}{\partial p_2}} \left(1 + \frac{\delta}{\mu}\right) \tau_1.$$

With $\tau_1 > 0$ to cover the fixed costs, and $\partial \tilde{x}_2^j / \partial p_2 < 0$, t_2 has the same sign as $\sum_j \pi^j \left(\partial \tilde{x}_2^j / \partial q_1\right) = \sum_j \pi^j \left(\partial \tilde{y}_1^j / \partial p_2\right)$. Thus if $\partial \tilde{y}_1^j / \partial p_2 > 0$, one sets $t_2 > 0$. This increases p_2 and, with it, $\tilde{y}_1^j$. On the other hand, if $\partial \tilde{y}_1^j / \partial p_2 < 0$, one sets $t_2 < 0$. This lowers p_2 and, as a result, increases $\tilde{y}_1^j$. Either way then, one sets t_2 to *increase* $\tilde{y}_1$. The break-even condition increases the price of y_1 above its marginal cost so that its consumption is less than optimal. One attempts to reverse this through t_2. *As a general lesson, taxation of privately-produced goods are necessitated to offset the distortions created by having to depart from marginal cost pricing on the part of the public firm.*[20]

To gain a better intuition into the nature of the tax/pricing rules in our model, we next consider the two well-known special cases for which the Ramsey setting yields simple results. These are the independent Hicksian and Marshallian demand curves cases whose solutions in the Ramsey model indicate the famous inverse elasticity rules (see (B9)–(B10) and (B11)–(B12) in Appendix B). Following these cases, we examine the most celebrated general result of the Ramsey model; namely the (un)equal proportional reduction in compensated demands property (see (B13)–(B14) in Appendix B). But first we summarize our main results thus far as,

Proposition 1 *Consider an Atkinson and Stiglitz setting wherein some goods (including the untaxed numeraire.) are produced by the private sector and some by a public or regulated firm subject to a break-even constraint. Then, contrary to the Atkinson and Stiglitz result,*

[20] The appearance of the δ/μ term on the right-hand side of this relationship reflects the fact that "average-cost pricing" by public firms creates an additional source of distortion beyond the tax distortion caused purely for revenue raising. Whereas the shadow cost of raising one unit of revenue for covering $\bar{R}$ is μ, it is $\mu + \delta$ for covering F. It is, relative to μ, δ/μ higher.

(i) Commodity taxes are desirable. Optimal commodity taxes are characterized by (15).

(ii) Break-even constraints for public/regulated firms have spill overs to other goods; they should be taxed too. These taxes are necessitated to offset the distortions created by having to depart from marginal cost pricing on the part of the public firm.

(iii) The tax/pricing rules for both types of goods, those that are produced privately and those that are provided through the public firm, are not affected by redistribution concerns.

4 Constrained Pareto efficient pricing rules

4.1 Zero cross-price compensated elasticities

Assume that Hicksian demands are independent so that the compensated demand of any produced good does not depend on the prices of the other produced goods. In this case, the reduced Slutsky matrix (where the line and column pertaining to leisure is deleted) is diagonal so that equation (15) simplifies to

$$\begin{pmatrix} t_2 \\ \vdots \\ t_n \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \\ \vdots \\ \left(1+\frac{\delta}{\mu}\right)\tau_m \end{pmatrix} = \frac{\delta}{\mu}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{\sum_j \pi^j y_1^j}{-\sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_1}} \\ \vdots \\ \frac{\sum_j \pi^j y_m^j}{-\sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial q_m}} \end{pmatrix}.$$

Consequently, for all $i = 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$\begin{aligned} t_i &= 0, \\ \tau_s &= \frac{\delta}{\mu+\delta}\frac{\sum_j \pi^j y_s^j}{-\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial q_s}} = \frac{\delta}{\mu+\delta}\frac{q_s \sum_j \pi^j y_s^j}{\sum_j \pi^j \tilde{y}_s^j \tilde{\varepsilon}_{ss}^j}, \end{aligned}$$

where $\widetilde{\varepsilon}_{ss}^j$ is the absolute value of the j-type's own-price elasticity of compensated demand for y_s. Or, for all $i = 2, \ldots, n$ and $s = 1, 2, \ldots, m$, $t_i = 0$ and

$$\frac{\tau_s}{1+\tau_s} = \frac{\delta \sum_j \pi^j y_s^j}{(\mu+\delta) \sum_j \pi^j y_s^j \widetilde{\varepsilon}_{ss}^j}, \tag{16}$$

which is an inverse elasticity rule. Again, the inverse elasticity rule arises only because of the existence of the break-even constraint. In the absence of this constraint, $\delta = 0$ so that τ_s, for all $s = 1, 2, \ldots, m$, will also be equal to zero.

The next question concerns the difference between our results of $\underline{t} = \underline{0}$ and $\underline{\tau}$ characterized by equation (16) with the inverse elasticity rules derived in the Ramsey framework (with a break-even constraint). The corresponding results under linear income taxes with independent Hicksian demands are reported in equations (B9)–(B10) in Appendix B. Comparing the two sets of results reveals two differences. One is that whereas in the traditional Ramsey model, both types of goods are subject to the inverse elasticity rules, in the Atkinson-Stiglitz framework only the goods that are produces by public firms are subject to the inverse elasticity rule. The goods produced by private firms should not be taxed. The second difference is that in the traditional Ramsey framework, the inverse elasticity rules are adjusted for redistributive concerns (through the covariance terms). No such terms appear in the Atkinson-Stiglitz framework. Here are redistributive concerns are taken care of by nonlinear income taxes. The only role for commodity taxes is "efficiency". There will be no tax on private goods while the goods provided by the public form are subject to the inverse elasticity rule that reflect pure efficiency considerations; see equation (16).

To sum, we find that in this special case, the private goods (not included in the break-even constraint) continue to go untaxed as in the Atkinson-Stiglitz setting with no break-even constraint. On the other hand, the pricing rules used by the public firm are purely efficiency-driven Ramsey rules. Goods are taxed inversely to their compensated demand elasticity regardless of their distributional implications. Redistribution is taken care of by the income tax allowing the public firm's prices to be adjusted for revenue raising (as in the Ramsey model with identical individuals).

It will become clear below that the apparent simplicity of this rule is to some extent misleading. It obscures some effects which are present but happen to cancel out in this special case. We shall return to this issue in the next subsections.

4.2 Zero cross-price elasticities

We now turn to the case where Marshallian demand functions are independent so that the demand for any given good does not depend on the prices of other (produced) goods.[21] To simplify the pricing rules that obtain in this case, it is simpler to start from the intermediate expressions (A10)–(A11) given in Appendix A rather than from (15). Rearranging these expressions, making use of the weak-separability assumption, and setting all the cross-price derivatives equal to zero, we obtain for all $i = 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$t_i \sum_j \pi^j \frac{\partial x_i^j}{\partial p_i} + \sum_j \pi^j x_i^j \left[\sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial c^j} + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial c^j} \right] = 0, \tag{17}$$

$$\left(1 + \frac{\delta}{\mu}\right) \tau_s \sum_j \pi^j \frac{\partial y_s^j}{\partial q_s} + \sum_j \pi^j y_s^j \left[\sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial c^j} + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial c^j} \right] + \frac{\delta}{\mu} \sum_j \pi^j y_s^j = 0. \tag{18}$$

Before simplifying these expressions any further, it is informative to delve into their interpretation. Recall that we are considering a compensated variation in the tax rates such that $dc_j = x_i^j dt_i$ for a variation in t_i and $dc_j = y_s^j dq_s$ for a variation in τ_s. In other words, individual disposable incomes are adjusted to keep utility levels constant for all individuals. With utility levels unchanged, the impact of the variation on social welfare entirely depends on the extra tax revenue (or profit) it generates. The left-hand sides

[21] Much of the regulation and industrial organization literature uses quasi-linear preference. In that case there are no income effects and the distinction between Hicksian and Marshallian demand becomes irrelevant.

of (17)–(18) measure the social value of this extra tax revenue (for a variation in t_i or in τ_s respectively). Obviously, when the tax system is optimized, this social value must be equal to zero (otherwise welfare could be increased by changing the tax rates).

To understand this interpretation, assume one changes c_j after t_i or τ_s changes. Start with a variation in t_i. With the tax revenues being given by

$$\sum_j \pi^j \left(\sum_{e=1}^{n} t_e x_e^j + \sum_{f=1}^{m} \tau_f y_f^j \right),$$

and the cross-price derivatives being equal to zero, the change in t_i produces an extra tax revenue of

$$\sum_j \pi^j x_i^j + t_i \sum_j \pi^j \frac{\partial x_i^j}{\partial p_i}.$$

Our compensation rule requires $\sum_j \pi^j x_i^j$ of this to be rebated to individuals.[22] The net change in revenue, minus compensation, is equal to

$$t_i \sum_j \pi^j \frac{\partial x_i^j}{\partial p_i}.$$

This is the first expression on the left-hand side of (17). At the same time, the $\sum_j \pi^j x_i^j$ compensation leads to an additional tax revenue of

$$\sum_{e=1}^{n} t_e x_e^j \left(\sum_j \pi^j x_i^j \right) + \sum_{f=1}^{m} \tau_f y_f^j \left(\sum_j \pi^j x_i^j \right) = \left(\sum_j \pi^j x_i^j \right) \left(\sum_{e=1}^{n} t_e x_e^j + \sum_{f=1}^{m} \tau_f y_f^j \right).$$

To convert these tax revenue changes into social welfare (measured in units of general revenues), one must multiply tax revenue variations emanating from $\underline{y}$-goods by a factor of $(1 + \delta/\mu)$. This is because the revenue from $\underline{y}$-goods enters both the global government budget constraint as well as the break-even constraint. This results in the second expression on the left-hand side of (17).

[22] To see this, observe that c_j changes according to $dc_j = x_i^j dt_i$ so that aggregate compensations change by $\sum_j \pi^j dc_j = \left(\sum_j \pi^j x_i^j \right) dt_i$.

The left-hand side of (18) can be decomposed in a similar way, except for one extra complication; namely the additional term $(\delta/\mu)\sum_j \pi^j y_s^j$. In this exercise, $\sum_j \pi^j y_s^j$ represents the value of the refunds to individuals. When collected as a tax, this amount has a social value of $(1+\delta/\mu)\sum_j \pi^j y_s^j$. On the other hand, the refund "costs" only $\sum_j \pi^j y_s^j$ (it comes from the general budget and has no impact on the break even constraint).

To ease the comparison with traditional Ramsey expressions, we can rewrite (17)–(18) as inverse elasticity rules. Introducing

$$A_i = \sum_{j=1}^{H} \pi^j x_i^j \left[\sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial c^j} + \left(1+\frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial c^j} \right], \tag{19}$$

$$B_s = \sum_{j=1}^{H} \pi^j y_s^j \left[\sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial c^j} + \left(1+\frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial c^j} \right] + \frac{\delta}{\mu} \sum_j \pi^j y_s^j, \tag{20}$$

where A_i and B_s measure the social value of the extra tax revenues due to refunds, with B_s also including $(\delta/\mu)\sum_j \pi^j y_s^j$. We have, for all $i = 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$\frac{t_i}{1+t_i} = \frac{A_i}{\sum_j \pi^j x_i^j \eta_{ii}^j}, \tag{21}$$

$$\frac{\tau_s}{1+\tau_s} = \frac{B_s}{(1+\delta/\mu)\sum_j \pi^j y_s^j \varepsilon_{ss}^j}, \tag{22}$$

where and η_{ii}^j and ε_{ss}^j denote the absolute value of the j-type's own-price elasticity of Marshallian demands for x_i and y_s.

Expressions (21)–(22) have a number of interesting implications, particularly when compared to their traditional counterparts. First, the effect of the break-even constraint is no longer confined to the goods which enter this constraint. Instead, it spills over to the other goods which no longer go untaxed (compare with the result obtained in Subsection 4.1). Second, we get inverse elasticity rules as in the Ramsey model; albeit without redistributive terms. This becomes clear below. The numerator of both expressions contain the "tax revenue" terms A_i and B_s. Recall that these expressions measure the social value of the extra tax revenue generated from the demand variations that follow the (compensating) adjustments in disposable income.

One may wonder why these terms were absent in Subsection 4.1. The key to understanding this property is that when Hicksian demands are independent, the price-cum-income variations we consider have by definition no impact on the demand of any of the other goods. And the effect on the good under consideration is already captured in the (compensated) elasticity term. To sum up, Subsection 4.1 has given simple results, not because the different effects were absent but because they happen to cancel out exactly under the considered assumptions.

Finally, compare (21)–(22) with their counterparts in the traditional Ramsey model under linear income taxes and independent Marshallian demands, as reported by (B11)–(B12) in Appendix B. The comparison reveals one significant difference. Our model results in the same exact expressions for optimal tax rates with one exception. Expressions (21)–(22) contain no terms reflecting redistributive concerns. These concerns enter in the Ramsey model through the covariance terms in (B11)–(B12).

4.3 Proportional reduction in compensated demands

When there are cross-price effects, the Ramsey model no longer yields results that can be presented as simple inverse elasticity rules. One popular way to present the tax rules in this case is in terms of proportional reduction in compensated demands. This leads to the celebrated "equal proportional reduction" pure efficiency result in the one-consumer Ramsey problem (and adjusted for redistributive considerations in the many household case).

$$-\frac{\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial p_e}\right) + \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial q_f}\right)}{\sum_j \pi^j x_i^j} = \frac{\delta}{\mu} \frac{\sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial p_i}\right)}{\sum_j \pi^j x_i^j}, \quad (23)$$

$$-\frac{\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial p_e}\right) + \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial q_f}\right)}{\sum_j \pi^j y_s^j} = \frac{\delta}{\mu} + \frac{\delta}{\mu} \frac{\sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial q_s}\right)}{\sum_j \pi^j y_s^j}. \quad (24)$$

The left-hand side of (23) and (24) represents the proportional reduction in compensated demand of x_i and y_i respectively. More precisely, it is proportional to the compensated impact of the considered good's tax rate on the break-even constraint. And as such, it differs across different commodities. Consequently, a version of the inverse elasticity rule also holds in the Atkinson-Stiglitz framework with a break-even constraint. However, the reduction is adjusted for tax revenue considerations; or more precisely the revenue of the regulated firm.

These are to be compared with (B13) and (B14) given in Appendix B for the Ramsey model. One again immediately observes that the two sets of formulas are identical except for the covariance terms that appear in (B13)–(B14). Equations (23) and (24) contain no such redistributive terms. Consequently, as in Subsection 4.2, nonlinear income taxation fully takes care of redistributive concerns and obviates the need to adjust the inverse elasticity rules for redistribution. We summarize the main results of this section as,

Proposition 2 *In the Atkinson and Stiglitz setting of Proposition 1:*

(i) Assume compensated demands are independent. Then, (a) the goods that are produced by non-regulated firms should not be taxed but public/regulated firms' goods should be. (b) Taxation of public/regulated firms' goods follow the Ramsey inverse elasticity rule as characterized by equation (16). (c) The taxation/pricing rules are purely efficiency-driven. Redistribution is taken care of by the income tax allowing the public

firm's prices to be adjusted for revenue raising (as in the Ramsey model with identical individuals).

(ii) Assume compensated demands are independent. Then, (a) All goods are taxed with the optimal taxes characterized by expressions (21)–(22). (b) Tax rates follows inverse elasticity rules as in the Ramsey model; albeit without redistributive terms. (c) The tax rules also include "tax revenue" terms characterized by (19)–(20). (d) These terms measure the social value of the extra tax revenue generated from the demand variations that follow the (compensating) adjustments in disposable income.

(iii) A version of the proportional reduction in compensated demands apply for all goods as characterized by (23) and (24). The reductions are proportional to the compensated impact of the considered good's tax rate on the break-even constraint and differs across different goods.

5 Conclusion

This paper has examined if the optimal tax and regulatory pricing approaches to Ramsey pricing can be reconciled. It has incorporated the two objectives of revenue raising for financing the government's expenditures (including redistributive transfers) and a regulated firm's fixed cost into a single framework. Tax instruments were restricted by informational considerations posited in the optimal tax literature and by the firm's break-even constraint as stipulated in the regulatory economics. It has argued that in practice regulation is almost always associated with balanced budget requirements. It is thus a shortcoming of the modern optimal tax theory that is based solely on a information structure that precludes break-even constraints. It has discussed a number of reasons for the existence of these constraints and shown that their neglect has led to an warranted view of the role of commodity taxes in the literature.

By incorporating a break-even constraint into the Atkinson and Stiglitz framework, the paper has challenged the modern optimal tax view that considers commodity taxes redundant. It has restored many of the earlier Ramsey tax/pricing lessons within the Atkinson and Stiglitz framework. In particular, it has shown that while nonlinear income taxes does take care of all redistributive concerns, this does not mean that commodity taxes are redundant. Break-even constraints create a role for commodity taxes. Interestingly too, this role goes beyond taxation of goods produced by the public/regulated firm but to the taxation of other goods as well. Put differently, break-even constraints for "regulated goods" have spill overs to "non-regulated goods".

We conclude by pointing out a number of possible extensions to our study. First, it would be interesting to compare the spill-over effects on the prices of non-regulated goods to the markups imposed on the goods subject to break-even constraints. Our various expressions suggest that this depends mainly on the size of the (compensated) cross-price effects. However, the complex way they operate does not allow one to draw clearcut conclusions at this level of generality. Numerical examples could provide some illustrative indications, while an empirical study may lead to more satisfactory answers.

Second, we have not formally modeled the informational asymmetries between public authorities and the regulated firm. These have been at the heart of the regulatory economics literature during the last two decades.[23] Such asymmetries of information would introduce an additional layer of complexity. For the regulator, the design of an incentive schemes come on top (and is intertwined) with the traditional price pricing problem; see *e.g.*, Laffont and Tirole (1993, chapter 3). However, the use of sophisticated incentive regulation does *not* in itself solve the problem of breaking even. In particular, this literature uses Ramsey pricing as a benchmark which obtains under full information. Under asymmetric information pertaining to a public/regulated firm's cost, pricing rules are more complex and often incorporate incentive corrections. As far as our results are

[23] As we have emphasized previously, our objective has been to explore what break-even constraints imply for the role of commodity taxes however one rationalizes these constraints. In a way, this is somewhat akin to the one-consumer Ramsey tax problem wherein one rules out lump-sum taxation for a variety of reasons and studies its implications for the design of optimal commodity taxes.

concerned, this complexity can only reinforce our main findings. The use of Ramsey-type prices in the Atkinson-Stiglitz setting, and the spillover to the taxation of goods produced in the private sector, are expected to be robust. What may not be robust, though, are the simple results we obtained for separable demands. However, even these results may well continue to hold, at least in some circumstances. In particular, Laffont and Tirole (1993, p. 173) show that when the cost function satisfies some separability conditions, one obtains the so-called "incentive-pricing dichotomy". Then, the incentive design leaves the pricing rules unaffected and we return to traditional Ramsey pricing.

Last but not least, regulation often pursues specific and likely non-welfarist redistributive objectives (as in universal access). It would be interesting to study how these objectives interact with the general objectives of tax policy.

Appendix A

First-order characterization of the(constrained) Pareto-efficient allocations: Rearranging the terms in (11), and dropping the constants $\bar{R}$ and F, one may usefully rewrite the Lagrangian expression as

$$\begin{aligned}\mathcal{L} \;=\; & \sum_j \left(\eta^j + \sum_{k \neq j} \lambda^{jk}\right) v^j + \mu \sum_j \pi^j \left[(I^j - c^j) + \sum_{i=1}^{n} (p_i - 1) x_i^j\right] \\ & + (\mu + \delta) \sum_j \pi^j \left[\sum_{s=1}^{m} (q_s - 1) y_s^j\right] - \sum_j \sum_{k \neq j} \lambda^{jk} v^{jk}. \qquad \text{(A1)}\end{aligned}$$

The first-order conditions of this problem are, for $j, k = 1, 2, \ldots, H$,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial I^j} = & \left(\eta^j + \sum_{k \neq j} \lambda^{jk}\right) v_I^j + \mu \pi^j \left[1 + \sum_{i=1}^{n} (p_i - 1) \frac{\partial x_i^j}{\partial I^j}\right] \\ & + (\mu + \delta)\, \pi^j \left[\sum_{s=1}^{m} (q_s - 1) \frac{\partial y_s^j}{\partial I^j}\right] - \sum_{k \neq j} \lambda^{kj} v_I^{kj} = 0, \qquad \text{(A2)} \\ \frac{\partial \mathcal{L}}{\partial c^j} = & \left(\eta^j + \sum_{k \neq j} \lambda^{jk}\right) v_c^j + \mu \pi^j \left[-1 + \sum_{i=1}^{n} (p_i - 1) \frac{\partial x_i^j}{\partial c^j}\right] \\ & + (\mu + \delta)\, \pi^j \left[\sum_{s=1}^{m} (q_s - 1) \frac{\partial y_s^j}{\partial c^j}\right] - \sum_{k \neq j} \lambda^{kj} v_c^{kj} = 0, \qquad \text{(A3)} \\ \frac{\partial \mathcal{L}}{\partial p_i} = & \sum_j \left(\eta^j + \sum_{k \neq j} \lambda^{jk}\right) v_i^j + \mu \sum_j \pi^j \left[\sum_{e=1}^{n} (p_e - 1) \frac{\partial x_e^j}{\partial p_i} + x_i^j\right] \\ & + (\mu + \delta) \sum_j \pi^j \left[\sum_{f=1}^{m} (q_f - 1) \frac{\partial y_f^j}{\partial p_i}\right] - \sum_j \sum_{k \neq j} \lambda^{jk} v_i^{jk} = 0, \quad i = 2, 3, \ldots, n, \qquad \text{(A4)}\end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial q_s} = \sum_j \left(\eta^j + \sum_{k \neq j} \lambda^{jk} \right) v_s^j + \mu \sum_j \pi^j \left[\sum_{e=1}^{n} (p_e - 1) \frac{\partial x_e^j}{\partial q_s} \right]$$
$$+ (\mu + \delta) \sum_j \pi^j \left[\sum_{f=1}^{m} (q_f - 1) \frac{\partial y_f^j}{\partial q_s} + y_s^j \right] - \sum_j \sum_{k \neq j} \lambda^{jk} v_s^{jk} = 0, \qquad s = 1, 2, \ldots, m, \tag{A5}$$

where a subscript on v^j denotes a partial derivative. Equations (A2)–(A5) characterize the Pareto-efficient allocations constrained both by the public firms' break-even constraint, the resource constraint, the self-selection constraints, as well as the linearity of commodity tax rates.

Derivation of (15) for optimal commodity taxes: Multiply equation (A3) by x_i^j, sum over j and add the resulting equation to (A4). Similarly, multiply (A3) by y_s^j, sum over j and add the resulting equation to (A5). Simplifying results in the following system of equations for $i = 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$\frac{\partial \mathcal{L}}{\partial p_i} + \sum_j x_i^j \frac{\partial \mathcal{L}}{\partial c^j} =$$
$$\sum_j \left(\eta^j + \sum_{k \neq j} \lambda^{jk} \right) \left(v_i^j + x_i^j v_c^j \right) + \mu \sum_j \pi^j \left[\sum_{e=1}^{n} (p_e - 1) \left(\frac{\partial x_e^j}{\partial p_i} + x_i^j \frac{\partial x_e^j}{\partial c^j} \right) \right]$$
$$+ (\mu + \delta) \sum_j \pi^j \left[\sum_{f=1}^{m} (q_f - 1) \left(\frac{\partial y_f^j}{\partial p_i} + x_i^j \frac{\partial y_f^j}{\partial c^j} \right) \right] - \sum_j \sum_{k \neq j} \lambda^{kj} \left(v_i^{kj} + x_i^j v_c^{kj} \right) = 0, \tag{A6}$$

$$\frac{\partial \mathcal{L}}{\partial q_s} + \sum_j y_s^j \frac{\partial \mathcal{L}}{\partial c^j} =$$
$$\sum_j \left(\eta^j + \sum_{k \neq j} \lambda^{jk} \right) \left(v_s^j + y_s^j v_c^j \right) + \mu \sum_j \pi^j \sum_{e=1}^{n} (p_e - 1) \left(\frac{\partial x_e^j}{\partial q_s} + y_s^j \frac{\partial x_e^j}{\partial c^j} \right)$$
$$+ (\mu + \delta) \sum_j \pi^j \left[\sum_{f=1}^{m} (q_f - 1) \left(\frac{\partial y_f^j}{\partial q_s} + y_s^j \frac{\partial y_f^j}{\partial c^j} \right) \right] - \sum_j \sum_{k \neq j} \lambda^{kj} \left(v_s^{kj} + y_s^j v_c^{kj} \right) + \delta \sum_j \pi^j y_s^j = 0. \tag{A7}$$

With $v_i^j + x_i^j v_c^j = 0$ from Roy's identity, the left-hand side of (A6) shows the impact on the Lagrangian expression $\mathcal{L}$ of a variation in p_i when the disposable income of individuals is adjusted according to

$$dc_j = x_i^j dt_i, \tag{A8}$$

to keep their utility levels constant. With $v_s^j + y_s^j v_c^j = 0$, the left-hand side of (A7) shows the same compensated effect for a variation in q_s where

$$dc_j = y_s^j d\tau_s. \tag{A9}$$

These compensated derivatives, $(\partial\mathcal{L}/\partial p_i)_{v^j=\overline{v}^j}$ and $(\partial\mathcal{L}/\partial q_s)_{v^j=\overline{v}^j}$ vanish at the optimal solution.

Make use of Roy's identity to set,

$$\begin{aligned} v_i^j + x_i^j v_c^j &= 0, \\ v_i^{kj} + x_i^{kj} v_c^{kj} &= 0, \\ v_s^j + y_s^j v_c^j &= 0, \\ v_s^{kj} + y_s^{kj} v_c^{kj} &= 0. \end{aligned}$$

Substitute these values in equations (A6)–(A7), set $p_i - 1 = t_i$ and $q_s - 1 = \tau_s$, and divide by μ. Upon changing the order of summation and further simplification one arrives at, for all $i = 2, \ldots, n$, and $s = 1, 2, \ldots, m$,

$$\begin{aligned} &\sum_{e=1}^{n} t_e \left[\sum_j \pi^j \left(\frac{\partial x_e^j}{\partial p_i} + x_i^j \frac{\partial x_e^j}{\partial c^j}\right)\right] + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left[\sum_j \pi^j \left(\frac{\partial y_f^j}{\partial p_i} + x_i^j \frac{\partial y_f^j}{\partial c^j}\right)\right] \\ &- \frac{1}{\mu} \sum_j \sum_{k \neq j} \lambda^{kj} \left(x_i^j - x_i^{kj}\right) v_c^{kj} = 0, \end{aligned} \tag{A10}$$

$$\begin{aligned} &\sum_{e=1}^{n} t_e \left[\sum_j \pi^j \left(\frac{\partial x_e^j}{\partial q_s} + y_s^j \frac{\partial x_e^j}{\partial c^j}\right)\right] + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left[\sum_j \pi^j \left(\frac{\partial y_f^j}{\partial q_s} + y_s^j \frac{\partial y_f^j}{\partial c^j}\right)\right] \\ &- \frac{1}{\mu} \sum_j \sum_{k \neq j} \lambda^{kj} \left(y_s^j - y_s^{kj}\right) v_c^{kj} + \frac{\delta}{\mu} \sum_j \pi^j y_s^j = 0. \end{aligned} \tag{A11}$$

Next, using the Slutsky equations,

$$\begin{aligned}
\frac{\partial x_e^j}{\partial p_i} &= \frac{\partial \tilde{x}_e^j}{\partial p_i} - x_i^j \frac{\partial x_e^j}{\partial c^j}, \\
\frac{\partial y_f^j}{\partial p_i} &= \frac{\partial \tilde{y}_f^j}{\partial p_i} - x_i^j \frac{\partial y_f^j}{\partial c^j}, \\
\frac{\partial x_e^j}{\partial q_s} &= \frac{\partial \tilde{x}_e^j}{\partial q_s} - y_s^j \frac{\partial x_e^j}{\partial c^j}, \\
\frac{\partial y_f^j}{\partial q_s} &= \frac{\partial \tilde{y}_f^j}{\partial q_s} - y_s^j \frac{\partial y_f^j}{\partial c^j},
\end{aligned}$$

while making use of the symmetry of the Slutsky matrix, one can further simplify (A10)–(A11) to

$$\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial p_e} \right) + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial q_f} \right) = \frac{1}{\mu} \sum_j \sum_{k \neq j} \lambda^{kj} \left(x_i^j - x_i^{kj} \right) v_c^{kj}, \tag{A12}$$

$$\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial p_e} \right) + \left(1 + \frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial q_f} \right) = \frac{1}{\mu} \sum_j \sum_{k \neq j} \lambda^{kj} \left(y_s^j - y_s^{kj} \right) v_c^{kj} - \frac{\delta}{\mu} \sum_j \pi^j y_s^j, \tag{A13}$$

which hold for all $i = 2, \ldots, n$, and $s = 1, 2, \ldots, m$. Finally, collect the terms involving δ/μ and use the definition of Δ in (14) to write out equations (A12)–(A13) in matrix notation:

$$\Delta \begin{pmatrix} t_2 \\ \vdots \\ t_n \\ \left(1 + \frac{\delta}{\mu}\right) \tau_1 \\ \vdots \\ \left(1 + \frac{\delta}{\mu}\right) \tau_m \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} \sum_j \sum_{k \neq j} \lambda^{kj} \left(x_2^j - x_2^{kj} \right) v_c^{kj} \\ \vdots \\ \sum_j \sum_{k \neq j} \lambda^{kj} \left(x_n^j - x_n^{kj} \right) v_c^{kj} \\ \sum_j \sum_{k \neq j} \lambda^{kj} \left(y_1^j - y_1^{kj} \right) v_c^{kj} \\ \vdots \\ \sum_j \sum_{k \neq j} \lambda^{kj} \left(y_m^j - y_m^{kj} \right) v_c^{kj} \end{pmatrix} - \frac{\delta}{\mu} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sum_j \pi^j y_1^j \\ \vdots \\ \sum_j \pi^j y_m^j \end{pmatrix}. \tag{A14}$$

Premultiplying through by Δ^{-1} yields

$$\begin{pmatrix} t_2 \\ \vdots \\ t_n \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \\ \vdots \\ \left(1+\frac{\delta}{\mu}\right)\tau_m \end{pmatrix} = \frac{1}{\mu}\Delta^{-1}\begin{pmatrix} \sum_j \sum_{k\neq j} \lambda^{kj}\left(x_2^j - x_2^{kj}\right) v_c^{kj} \\ \vdots \\ \sum_j \sum_{k\neq j} \lambda^{kj}\left(x_n^j - x_n^{kj}\right) v_c^{kj} \\ \sum_j \sum_{k\neq j} \lambda^{kj}\left(y_1^j - y_1^{kj}\right) v_c^{kj} \\ \vdots \\ \sum_j \sum_{k\neq j} \lambda^{kj}\left(y_m^j - y_m^{kj}\right) v_c^{kj} \end{pmatrix} - \frac{\delta}{\mu}\Delta^{-1}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sum_j \pi^j y_1^j \\ \vdots \\ \sum_j \pi^j y_m^j \end{pmatrix}. \tag{A15}$$

With weak-separability of preferences, equations (12)–(13) hold so that $x_i^{jk} = x_i^k$ and $y_s^{jk} = y_s^k$. It then follows immediately that the first vector on the right-hand side of (A15) vanishes, reducing it to (15).

Derivation of (23)–(24): With weakly-separable preferences, one can rearrange equations (A12)–(A13) as

$$\begin{aligned} -\mu\sum_{e=1}^{n} t_e\left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial p_e}\right) - \mu\sum_{f=1}^{m}\tau_f\left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial q_f}\right) &= \delta\sum_{f=1}^{m}\tau_f\left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial p_i}\right), \\ -\mu\sum_{e=1}^{n} t_e\left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial p_e}\right) - \mu\sum_{f=1}^{m}\tau_f\left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial q_f}\right) &= \delta\sum_{f=1}^{m}\tau_f\left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial q_s}\right) + \delta\sum_j \pi^j y_s^j. \end{aligned}$$

Then divide the first set of equations by $\mu\sum_j \pi^j x_i^j$ and the second set of equations by $\mu\sum_j \pi^j y_s^j$. They will then be rewritten as (23)–(24).

Appendix B: The benchmark

The many-consumer Ramsey problem with a break-even constraint

An individual of type j now chooses $\underline{x}, \underline{y}$, and L to maximize his utility $u = u\left(\underline{x}, \underline{y}, L\right)$ subject to the budget constraint $\sum_i p_i x_i + \sum_s q_s y_s = w^j (1-\theta) L + b$, where θ is the linear income tax rate and b is the uniform lump-sum rebate. The resulting demand functions for $\underline{x}$ and $\underline{y}$, and the supply function for L, for all $i = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, m$, are denoted by $x_i^j = x_i\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right), y_s^j = y_s\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right)$, and $L^j = L\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right)$. Substituting in the utility function yields the indirect utility function[24]

$$\begin{aligned} v^j &= v\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right) \\ &\equiv u\left(\underline{x}\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right), \underline{y}\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right), L\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right)\right). \end{aligned}$$

Let π^j denote the proportion of individuals of type j in the economy and consider a utilitarian social welfare function of the form

$$\sum_{j=1}^{H} \pi^j W\left(v^j\right),$$

where $W(\cdot)$ is increasing, twice differentiable, and concave. The optimization problem consists of maximizing $\sum_j \pi^j W\left(v^j\right)$ with respect to $\underline{p}, \underline{q}, b$ and subject to resource and break-even constraints (8)–(9). Observe that with $\underline{x}, \underline{y}$ and L being homogeneous of

[24] Separability of preferences does not simplify the characterization of the *unconditional* demand and indirect utility functions. If the utility function is separable and is written as $U\left(f\left(\underline{x}, \underline{y}\right), L\right)$, all one can do is to write, *for a given* L, the consumer's demand for $\underline{x}, \underline{y}$ as functions of $\underline{p}, \underline{q}$ and $w^j (1-\theta) L + b$. Then,

$$v^j = U\left(f\left(\underline{x}\left(\underline{p}, \underline{q}, w^j (1-\theta) L + b\right), \underline{y}\left(\underline{p}, \underline{q}, w^j (1-\theta) L + b\right)\right), L\right).$$

However, optimization over L yields $L = L\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right)$. Consequently,

$$v^j = v\left(\underline{p}, \underline{q}, w^j (1-\theta), b\right).$$

degree zero in $\underline{p}, \underline{q}, w^j(1-\theta)$ and b, consumer prices are determined only up to a proportionality factor. It is for this reason that we do not optimize over θ, setting it equal to zero.

The Lagrangian expression associated with maximizing $\sum_j \pi^j W\left(v^j\right)$ with respect to $\underline{p}, \underline{q}, b$ and subject to constraints (8)–(9) is represented by

$$\begin{aligned}
\mathcal{L} \;=\; & \sum_j \pi^j W\left(v^j\right) + \mu \left\{ \sum_j \pi^j \left[\sum_{i=1}^{n} (p_i - 1) x_i^j + \sum_{s=1}^{m} (q_s - 1) y_s^j - b \right] - \bar{R} - F \right\} \\
& + \delta \left\{ \sum_j \pi^j \left[\sum_{s=1}^{m} (q_s - 1) y_s^j \right] - F \right\}.
\end{aligned}$$

The first-order conditions of this problem are,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial b} = & \sum_j \pi^j \frac{\partial W}{\partial v^j} \frac{\partial v^j}{\partial b} + \mu \sum_j \pi^j \left[-1 + \sum_{i=1}^{n} (p_i - 1) \frac{\partial x_i^j}{\partial b} \right] + \\
& (\mu + \delta) \sum_j \pi^j \left[\sum_{s=1}^{m} (q_s - 1) \frac{\partial y_s^j}{\partial b} \right] = 0, \qquad \text{(B1)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial p_i} = & \sum_j \pi^j \frac{\partial W}{\partial v^j} \frac{\partial v^j}{\partial p_i} + \mu \sum_j \pi^j \left[x_i^j + \sum_{e=1}^{n} (p_e - 1) \frac{\partial x_e^j}{\partial p_i} \right] + \\
& (\mu + \delta) \sum_j \pi^j \left[\sum_{f=1}^{m} (q_f - 1) \frac{\partial y_f^j}{\partial p_i} \right] = 0, \quad i = 1, 2, \ldots, n, \qquad \text{(B2)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial q_s} = & \sum_j \pi^j \frac{\partial W}{\partial v^j} \frac{\partial v^j}{\partial q_s} + \mu \sum_j \pi^j \left[\sum_{e=1}^{n} (p_e - 1) \frac{\partial x_e^j}{\partial q_s} \right] + \\
& (\mu + \delta) \sum_j \pi^j \left[y_s^j + \sum_{f=1}^{m} (q_f - 1) \frac{\partial y_f^j}{\partial q_s} \right] = 0, \quad s = 1, 2, \ldots, m. \qquad \text{(B3)}
\end{aligned}$$

Define

$$\begin{aligned}
\alpha^j &\equiv \frac{\partial v^j}{\partial b}, \\
\beta^j &\equiv \frac{1}{\mu}\frac{\partial W}{\partial v^j}\frac{\partial v^j}{\partial b}, \\
\gamma^j &\equiv \beta^j + \sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial b} + \left(1+\frac{\delta}{\mu}\right)\sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial b}.
\end{aligned}$$

Using these definitions, the first-order condition with respect to b, equation (B1), is simplified to

$$\sum_j \pi^j \gamma^j = 1.$$

Then using Roy's identity, the first-order conditions (B2)–(B3) are simplified to, for all $i = 1, 2, \ldots, n$, and $s = 1, 2, \ldots, m$,

$$\sum_j \pi^j \left[-\beta^j x_i^j + \left(x_i^j + \sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial p_i} \right) + \left(1+\frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial p_i} \right] = 0, \tag{B4}$$

$$\sum_j \pi^j \left[-\beta^j y_s^j + \left(\sum_{e=1}^{n} t_e \frac{\partial x_e^j}{\partial q_s} \right) + \left(1+\frac{\delta}{\mu}\right) \left(y_s^j + \sum_{f=1}^{m} \tau_f \frac{\partial y_f^j}{\partial q_s} \right) \right] = 0. \tag{B5}$$

Next using the Slutsky equation and the symmetry of the Slutsky matrix, equations (B4)–(B5) can also be written as

$$-\left[\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \widetilde{x}_i^j}{\partial p_e} \right) + \left(1+\frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \widetilde{x}_i^j}{\partial q_f} \right) \right] = \sum_j \pi^j x_i^j - \sum_j \gamma^j \pi^j x_i^j, \tag{B6}$$

$$-\left[\sum_{e=1}^{n} t_e \left(\sum_j \pi^j \frac{\partial \widetilde{y}_s^j}{\partial p_e} \right) + \left(1+\frac{\delta}{\mu}\right) \sum_{f=1}^{m} \tau_f \left(\sum_j \pi^j \frac{\partial \widetilde{y}_s^j}{\partial q_f} \right) \right] = \left(1+\frac{\delta}{\mu}\right) \sum_j \pi^j y_s^j - \sum_j \gamma^j \pi^j y_s^j. \tag{B7}$$

Finally, define

$$\Omega \equiv \begin{pmatrix} \sum_j \pi^j \frac{\partial \tilde{x}_1^j}{\partial p_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_1^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{x}_1^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_1^j}{\partial q_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial p_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{x}_n^j}{\partial q_m} \\ \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial p_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_1^j}{\partial q_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial p_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial p_n} & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial q_1} & \cdots & \sum_j \pi^j \frac{\partial \tilde{y}_m^j}{\partial q_m} \end{pmatrix}$$

to rewrite equations (B6)–(B7) in matrix notation

$$\Omega \begin{pmatrix} t_1 \\ \vdots \\ t_n \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \\ \vdots \\ \left(1+\frac{\delta}{\mu}\right)\tau_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^H \left(1-\gamma^j\right)\pi^j x_1^j \\ \vdots \\ \sum_{j=1}^H \left(1-\gamma^j\right)\pi^j x_n^j \\ \sum_{j=1}^H \left(1+\frac{\delta}{\mu}-\gamma^j\right)\pi^j y_1^j \\ \vdots \\ \sum_{j=1}^H \left(1+\frac{\delta}{\mu}-\gamma^j\right)\pi^j y_m^j \end{pmatrix}.$$

Pre-multiplying by Ω^{-1} yields,[25]

$$\begin{pmatrix} t_1 \\ \vdots \\ t_n \\ \left(1+\frac{\delta}{\mu}\right)\tau_1 \\ \vdots \\ \left(1+\frac{\delta}{\mu}\right)\tau_m \end{pmatrix} = \Omega^{-1} \begin{pmatrix} \sum_{j=1}^H \left(1-\gamma^j\right)\pi^j x_1^j \\ \vdots \\ \sum_{j=1}^H \left(1-\gamma^j\right)\pi^j x_n^j \\ \sum_{j=1}^H \left(1+\frac{\delta}{\mu}-\gamma^j\right)\pi^j y_1^j \\ \vdots \\ \sum_{j=1}^H \left(1+\frac{\delta}{\mu}-\gamma^j\right)\pi^j y_m^j \end{pmatrix}. \tag{B8}$$

Observe that while the structure of taxes on private goods and public firms' goods are not identical here, the same Ramsey tax/pricing rules apply to both under this setting (as compared to our result in the text given by equation (15)). Observe also that the many-consumer Ramsey problem we have considered as our benchmark here allows for a uniform lump-sum rebate, b. The literature considers this problem for both

[25] As with Δ, Ω is of full rank so that its inverse exists; see Takayama (1985).

cases when b is and is not present. The same tax/pricing rules are derived in both cases. The only difference is that when b is present, the optimization over b results in $\sum_j \pi^j \gamma^j = 1$. This result does not hold when b is not present.

Zero cross-price compensated elasticities

Set the cross-price derivatives in equations (B6)–(B7) equal to zero and rearrange the terms. We have, for all $i = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$t_i = \frac{\sum_j \left(\gamma^j - 1\right) \pi^j x_i^j}{\sum_j \pi^j \frac{\partial \widetilde{x}_i^j}{\partial p_i}},$$

$$\tau_s = \frac{\sum_j \left(\gamma^j - 1\right) \pi^j y_s^j - \frac{\delta}{\mu} \sum_j \pi^j y_s^j}{\left(1 + \frac{\delta}{\mu}\right) \sum_j \pi^j \frac{\partial \widetilde{y}_s^j}{\partial q_s}}.$$

Denoting the absolute value of the j-type's own-price elasticity of compensated demands for x_i and y_s by $\widetilde{\eta}_{ii}^j$ and $\widetilde{\varepsilon}_{ss}^j$, one can rewrite the above expressions as,

$$\frac{t_i}{1 + t_i} = \frac{\sum_j \left(1 - \gamma^j\right) \pi^j x_i^j}{\sum_j \pi^j x_i^j \widetilde{\eta}_{ii}^j} = \frac{-\mathrm{Cov}\left(x_i^j, \gamma^j\right)}{\sum_j \pi^j x_i^j \widetilde{\eta}_{ii}^j}, \tag{B9}$$

$$\frac{\tau_s}{1 + \tau_s} = \frac{\sum_j \left(1 + \delta/\mu - \gamma^j\right) \pi^j y_s^j}{\left(1 + \delta/\mu\right) \sum_j \pi^j y_s^j \widetilde{\varepsilon}_{ss}^j} = \frac{\delta \sum_j \pi^j y_s^j - \mathrm{Cov}\left(y_s^j, \gamma^j\right)}{\left(\mu + \delta\right) \sum_j \pi^j y_s^j \widetilde{\varepsilon}_{ss}^j}, \tag{B10}$$

where Cov(.,.) denotes covariance and the covariance interpretation follows because of the result that $\sum_j \gamma^j = 1$.

Zero cross-price elasticities

Set the cross-price derivatives in equations (B4)–(B5) equal to zero and rearrange the terms to get

$$\frac{t_i}{1 + t_i} = \frac{\sum_j \left(1 - \beta^j\right) \pi^j x_i^j}{\sum_j \pi^j x_i^j \eta_{ii}^j},$$

$$\frac{\tau_s}{1 + \tau_s} = \frac{\sum_j \left(1 + \delta/\mu - \beta^j\right) \pi^j y_s^j}{\left(1 + \delta/\mu\right) \sum_j \pi^j y_s^j \varepsilon_{ss}^j}.$$

Add and subtract γ^j to β^j to rewrite the above expressions as

$$\begin{aligned}
\frac{t_i}{1+t_i} &= \frac{\sum_j \left(\gamma^j - \beta^j\right)\pi^j x_i^j - \sum_j \left(\gamma^j - 1\right)\pi^j x_i^j}{\sum_j \pi^j x_i^j \eta_{ii}^j}, \\
\frac{\tau_s}{1+\tau_s} &= \frac{\sum_j \left(\gamma^j - \beta^j\right)\pi^j y_s^j + (\delta/\mu)\sum_j \pi^j y_s^j - \sum_j \left(\gamma^j - 1\right)\pi^j y_s^j}{(1+\delta/\mu)\sum_j \pi^j y_s^j \varepsilon_{ss}^j}.
\end{aligned}$$

Finally, rewrite $\sum_j \left(\gamma^j - 1\right)\pi^j x_i^j$ and $\sum_j \left(\gamma^j - 1\right)\pi^j y_s^j$ as $\mathrm{Cov}\left(x_i^j, \gamma^j\right)$ and $\mathrm{Cov}\left(y_s^j, \gamma^j\right)$. Then recall from the definition of γ^j that we have: $\gamma^j - \beta^j = \sum_{e=1}^n t_e \left(\partial x_e^j/\partial b\right) + (1+\delta/\mu)\sum_{f=1}^m \tau_f \left(\partial y_f^j/\partial b\right)$. Substitute in above and make use of the definitions of A_i and B_s in the text to arrive at the counterparts of (21)–(22) in the text. We have, for all $i = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, m$,

$$\frac{t_i}{1+t_i} = \frac{A_i - \mathrm{Cov}\left(x_i^j, \gamma^j\right)}{\sum_j \pi^j x_i^j \eta_{ii}^j}, \tag{B11}$$

$$\frac{\tau_s}{1+\tau_s} = \frac{B_s - \mathrm{Cov}\left(y_s^j, \gamma^j\right)}{(1+\delta/\mu)\sum_j \pi^j y_s^j \varepsilon_{ss}^j}. \tag{B12}$$

Proportional reduction in compensated demands

Rearrange equations (B6)–(B7) and substitute $-\mathrm{Cov}\left(x_i^j, \gamma^j\right)$ for $\sum_j \left(1 - \gamma^j\right)\pi^j x_i^j$ and $-\mathrm{Cov}\left(y_s^j, \gamma^j\right)$ for $\sum_j \left(1 - \gamma^j\right)\pi^j y_s^j$ in them. We get, as counterparts of equations (23)–(24) in the text,

$$-\frac{\sum_{e=1}^n t_e \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial p_e}\right) + \sum_{f=1}^m \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{x}_i^j}{\partial q_f}\right)}{\sum_j \pi^j x_i^j} = \frac{\frac{\delta}{\mu}\sum_{f=1}^m \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial p_i}\right) - \mathrm{Cov}\left(x_i^j, \gamma^j\right)}{\sum_j \pi^j x_i^j}, \tag{B13}$$

$$-\frac{\sum_{e=1}^n t_e \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial p_e}\right) + \sum_{f=1}^m \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_s^j}{\partial q_f}\right)}{\sum_j \pi^j y_s^j} = \frac{\delta}{\mu} + \frac{\frac{\delta}{\mu}\sum_{f=1}^m \tau_f \left(\sum_j \pi^j \frac{\partial \tilde{y}_f^j}{\partial q_s}\right) - \mathrm{Cov}\left(y_s^j, \gamma^j\right)}{\sum_j \pi^j y_s^j}. \tag{B14}$$

References

[1] Allais, Maurice, 1947. Le problème de la coordination des transports et la théorie economique, *Bulletin des Ponts et Chaussées et des Mines.*

[2] Atkinson, Anthony B. and Joseph E. Stiglitz, 1976. The design of tax structure: direct versus indirect taxation, *Journal of Public Economics*, **6**, 55–75.

[3] Baumol, William J. and David F. Bradford, 1970. Optimal departures from marginal cost pricing, *American Economic Review*, **60**, 265–283.

[4] Boiteux, Marcel, 1956. Sur la gestion des monopoles publics astreints à l'équilibre budgétaire, *Econometrica,* **24**, 22–40. Published in English as: On the management of public monopolies subject to budgetary constraints, *Journal of Economic Theory*, **3**, 219–240, 1971.

[5] Brown, Stephen and David Sibley, 1986. *The Theory of Public Utility Pricing,* Cambridge University Press, Cambridge, UK.

[6] Christiansen, Vidar, 1984. Which commodity taxes should supplement the income tax? *Journal of Public Economics,* **24***,* 195–220.

[7] Coase, Ronald, 1945. Price and output policy of state enterprise, *Economics Journal,* **55**, 112–113.

[8] Coase, Ronald, 1946. The marginal cost controversy, *Economica*, **13,** 169–182

[9] Cremer, Helmuth and Firouz Gahvari, 1995. Uncertainty, optimal taxation and the direct versus indirect tax controversy, *Economic Journal*, **105**, 1165–79.

[10] Cremer, Helmuth and Firouz Gahvari, 1997. In–kind transfers, self–selection and optimal tax policy. *European Economic Review*, **41**, 97–114.

[11] Cremer, Helmuth and Firouz Gahvari, 1998. On optimal taxation of housing", *Journal of Urban Economics*, **43**, 315–335.

[12] Cremer, Helmuth and Firouz Gahvari, 2002. Nonlinear pricing, redistribution and optimal tax policy, *Journal of Public Economic Theory,* **4**, 139–161.

[13] Cremer, Helmuth, Firouz Gahvari and Norbert Ladoux, 1998. Externalities and optimal taxation, *Journal of Public Economics*, **70**, 343–364.

[14] Cremer, Helmuth, Fahrid Gasmi, André Grimaux and Jean-Jacqaues Laffont, 2001. Universal Service: an economic perspective, *Annals of Public and Cooperative Economics*, **72**, 5–43.

[15] Cremer, Helmuth, Pierre Pestieau and Jean-Charles Rochet, 2001. Direct versus indirect taxation : The design of the tax structure revisited, *International Economic Review*, **42**, 781–799.

[16] Crew, Michael and Paul Kleindorfer, 2011. Competitive strategies under FMO and intermodal competition, in *Reinventing the Postal Sector in an Electronic Age*, M. A. Crew and P. R. Kleindorfer, ed., Northampton: Edward Elgar Publishing, Inc., 1–19.

[17] Crew, Michael and Paul Kleindorfer, 2012. Non-linear pricing and volume discounts and the USO under entry, in *Multi-modal Competition and the Future of Mail*, M. A. Crew and P. R. Kleindorfer, ed., Northampton: Edward Elgar Publishing, Inc., 1–18.

[18] Diamond, Peter A. and James A. Mirrlees, 1971. Optimal taxation and public production I–II, *American Economic Review*, **61**, 8–27, 261–278.

[19] Diamond, Peter A., 1975. A many-person Ramsey tax rule, *Journal of Public Economics,* **4,** 335–342.

[20] Laffont, Jean-Jacques and Jean Tirole, 1990. Bypass and cream skimming, *American Economic Review,* **80**, 1042–1061.

[21] Laffont, Jean-Jacques and Jean Tirole, 1993. A Theory of Incentives in Procurement and Regulation, MIT Press, Cambridge, Mass.

[22] Laffont, Jean-Jacques and Jean Tirole, 2000. Competition in Telecommunications, MIT Press, Cambridge, Mass.

[23] Mirrlees, James A. 1971. An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*, **38**, 175–208.

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

[25] Sandmo, Agnar, 1976. Optimal taxation: An introduction to the literature, *Journal of Public Economics,* **6,** 37–54.

[26] Stiglitz, Joseph E., 2015. In praise of Frank Ramsey's contribution to the theory of taxation. *Economic Journal,* **125**, 235–268.

[27] Takayama, Akira. 1985. *Mathematical Economics*, 2nd ed. Cambridge: Cambridge University Press.

[28] Viscusi, W. Kip, John Vernon and Joseph Harrington, 1998. *Economics of Regulation and Antitrust*, 2nd ed. MIT Press, Cambridge, MA.

[29] Vogelsang, Ingo et Jörg Finsinger, 1979. Regulatory adjustment process for optimal pricing of multiproduct firm, *Bell Journal of Economics*, **10**, 157–171.

[30] Ware, Roger and Ralph A. Winter, 1986. Public pricing under imperfect competition, *International Journal of Industrial Organization*, **4**, 87–97.