Discussion Papers in Economics

Some principles for corrective taxation of externalities in a second-best world with commodity taxes

Sushama Murty
Discussion Paper 18-01

Centre for International Trade and Development

School of International Studies

Jawaharlal Nehru University

India

*

Some principles for corrective taxation of externalities in a second-best world with commodity taxes

Sushama Murty*

January 2, 2018

Abstract

Some principles for the design of commodity-taxes are derived, when they are employed for correcting externalities in addition to meeting the conventional revenue and redistributive objectives of the government. With production externalities, *production efficiency* is violated at the second-best, which involves intermediate-input taxation of (i)the externality-causing good directly, *or* (ii)"non-substitutable inputs" or outputs produced by them in manufacturing chains that include the externality-generating goods, *or* (iii)commodities whose use as inputs is complementary to the input usage of goods that generate externalities, *or* (iv)some combinations of (i)–(iii). Second-best consumption taxes have two independent and additive components: (a)a conventional equity and efficiency-balancing "many person Ramsey rule"(MPRR)-based VAT or GST and (b)an "externality-correcting" excise duty. The externality components of both consumption and optimal intermediate-input taxes are linked and cannot be chosen independently, although there exist several *degrees of freedom* in selecting them at a second-best. The optimal VAT is zero for commodities with indirectly-derived demands such as electricity, motoring-fuel, and road-services. The input-tax credit to a producer, who pays retail price for an intermediate-input, is equal to the GST plus the excess of the externality-excise component of the consumption tax over the corresponding intermediate-input tax. Many real-life commodity-tax policies are compared with our results.

Keywords: *optimal commodity taxation, production efficiency, production and consumption externalities, many person Ramsey rule, non-substitutable inputs, the value added tax, excise taxes, intermediate-input taxes*
JEL classification codes: *H21, H23, D51*

*Centre for International Trade and Development, School of International Studies, Jawaharlal Nehru University, New Delhi 110067.
Email: sushama.murty1@gmail.com

Some principles for corrective taxation of externalities in a second-best world.

by

Sushama Murty

1 Introduction

In two seminal articles, Diamond and Mirrlees (1971 a,b) (henceforth, DM) resurrected Ramsey's (1927) work on optimal commodity taxation and demonstrated that, when commodity taxes have to be employed in lieu of personalised lump-sum transfers as second-best instruments of revenue generation and redistribution by the government, then the optimal commodity tax structure satisfies the many person Ramsey rule (MPRR) and results in an allocation that is production efficient.

Balancing the equity and efficiency considerations, the MPRR recommends a system of differential taxation of commodities depending on (i) how their total consumptions are distributed between different consumers, (ii) the society's marginal valuation of welfare levels and abilities to pay of different (*e.g.,* rich and poor) consumers, and (iii) the differential responsiveness of tax payments of consumers to tax-induced changes in their real incomes. Production efficiency, on the other hand, is a feature of first-best Pareto optimal allocations that is also retained by second-best optimal allocations attained under commodity taxation. There are three key assumptions in the DM framework that underly this result, the first two of which are technological constant returns to scale and local Pareto non-satiation.[1] Extending the Diamond and Mirrlees framework to incorporate technological non-increasing returns, Dasgupta and Stiglitz (1972) demonstrated that production efficiency continues to be desirable in second-best economies with commodity taxation if the government can also implement a system of firm-specific profit taxation. A more general result is obtained in Murty (2013), which distinguishes between the profit taxation power of the government and the institutional rules of distribution of profits to consumers in the economy. It is shown that second-best production efficiency is desirable when the two are well-aligned.

The DM theoretical results characterising an optimal system of commodity taxation have greatly influenced the design of modern-day commodity tax structures such as the value added tax (VAT) or the goods and services tax (GST) and the retail sales tax (RST). While administrative and monitoring constraints restrict the number of commodity tax rates that can practically be implemented under these systems, it is desirable that even the few rates that end up being implemented should, in principle, be chosen keeping in mind both equity and efficiency considerations as noted by the MPRR.[2] Modern-day tax structures are also designed

[1]Intuitively, local Pareto non-satiation holds when the government has access to policy instruments, such as a uniform lump-sum transfer or a commodity tax on a commodity that is universally liked, which it can use to increase welfare of all consumers whenever it finds itself in a situation where its budget is in surplus.

[2]Tax policies in many OECD countries are converging to tax systems with very few VAT rates, if not a

to incorporate the principle of production efficiency, *e.g.,* under the VAT, taxes paid by firms on their purchases of inputs are credited back to them, while the RST is assessed only on transactions between firms and consumers.

The third key assumption of the DM(a) analysis is the absence of externalities. It is well-known that, even in first-best worlds, conditions for Pareto optimality are modified in the presence of non-rival (public good-type) externalities.[3] Implementation of a Pareto optimum in a market economy with the help of fiscal instruments requires Pigou (commodity) taxation/ subsidisation of the externality *in addition to* personalised lump-sum transfers. When the latter instruments are not available, Ramsey commodity taxation can step in to serve all three purposes of governmental revenue generation, redistribution, and externality correction.

There is a considerable theoretical literature that studies externality policies in second-best frameworks. Works such as Sandmo (1975) and those in the literature on double dividends,[4] study externalities generated by consumers that affect the well-being of consumers in a world where all consumers are identical and technologies exhibit constant returns to scale. As discussed in DM, in such models with consumption externalities, the second-best optimal commodity tax structure continues to exhibit production efficiency; so there continues to be no role for intermediate input taxation at a second-best optimum. The main result of this literature is that the second-best optimal tax on a consumption externality-causing good is a sum of a Ramsey component whose role is to minimise the conventional dead-weight loss due to commodity taxation and a Pigouvian tax whose role is to regulate the consumption externality.

Some important works that study the extensions of DM and Sandmo (1975) to the case of production externalities are Mayeres and Proost (1994, 1997) (henceforth, MP), Bovenberg and Goulder (1996), and De Borger (1996). The literature has been further surveyed with a special focus on the transport sector by Ahlberg (2006). These works show that inter-firm transactions in intermediate-inputs that cause externalities need to be taxed at a second-best optimum, and that such taxes are purely Pigouvian in nature.

It is well-known that in the presence of reciprocal externalities, the appropriate definition of a market equilibrium is a Nash equilibrium. In first-best worlds with externalities, Pigouvian taxation of externalities along with personalised lump-sum transfers helps achieve any Pareto optimum as a Nash equilibrium, with the optimal Pigouvian taxes ensuring that the generators will generate the Pareto-optimal levels of externalities, whose amounts will be taken as given by the victims. However, in much of the second-best literature on externalities that has been

uniform rate of VAT, based on a set of theoretical results (see *e.g.,* Atkinson and Stiglitz (1976, 1980) and Mirrlees (1976)) that demonstrated the superiority of non-linear income tax over linear commodity taxation as a redistributive instrument. However, these results rest on the assumption that consumer preferences are separable in consumption goods and leisure/labour. Moreover, they appear to be less relevant for a significant proportion of the world characterised by weak institutions, a huge informal sector including agriculture that are out of the ambit of income taxation, and large black economies. In such situations, the incentive compatibility of the Mirrleesian income tax is jeopardised and, till such institutional constraints are removed, commodity taxes will continue to play an important role in meeting government's redistributive objectives. See Murty (2017).

[3]See for example, Baumol and Oates (1988).

[4]See, *e.g.,* Bovenberg and van der Ploeg (1994), Bovenberg and de Mooij (1994), Goulder (1995), and Bovenberg and Goulder (1996).

surveyed above, the definitions of tax equilibria employed do not imply Nash equilibria in the externality-dimensions. As a result, while computing the net social marginal costs of externalities, this literature *either* misses out on the effects of changes in the levels of externalities on the tax-revenue (as, *e.g.,* in Sandmo (1975) and Bovenberg and Goulder (1996)) that arise due to the complementarities and substitutabilities between externality-causing goods and other goods[5] *or* includes some *non*-Nash like feedback effects of commodity-tax induced changes in externalities on consumer demands (as, *e.g.,* in MP and De Borger (1996)).

In contrast to the rest of the externality literature surveyed above, which assumes that all consumers are identical, MP and De Borger (1996) address the redistributary role of commodity taxes by studying economies with heterogeneous agents. Thus, their analysis extends the DM many-person economy analysis to the case of production and consumption externalities. MP argue that, in this case, the second-best optimal tax structure is similar to the one derived by Sandmo (1975) for the identical consumers case, in that it is additive in nature; being composed of (i) an externality component (which they call the net social Pigouvian tax) that only enters the tax formulae for only the externality-causing goods and (ii) the conventional MPRR component that enters the tax formulae of all goods. It is popularly known that the MPRR can be expressed in terms of normalised covariances between consumption of goods by consumers and the net social marginal utility of income to consumers.[6] In the analyses of MP, as opposed to DM and Atkinson and Sitglitz (1976, 1981), the net social marginal utility of income to a consumer has an additional component which also depends on the net social Pigouvian tax.[7] Hence, for all goods, the MPRR components of the tax formulae given in MP also depend on the net social Pigouvian tax, which seems contradictory to their claim that the externality component enters only the tax formulae for the externality-causing goods. Thus, in their work, both the MPRR and externality components of the optimal commodity tax rates depend on the net social Pigouvian tax.

In addition to the assumptions made by MP, De Borger (1996) allows technological non-increasing returns, with profits of firms being distributed to consumers in proportion to the shares that they own in the firms. However, as his model excludes profit taxation, it follows from Dasgupta and Stiglitz (1972) and Murty (2013) that, in De Borger's model, minimisation of dead-weight losses due to commodity taxation may itself require taxation of inter-firm transactions in intermediate inputs, quite apart from the prevalence of production externalities.

An influential body of less-theoretical policy-oriented literature[8] that reviews and assesses real-life policies of the government and informs and makes recommendations to it argues heuristically that the government also has the option (depending, for example, on its administrative convenience) of taxing units located upstream or downstream in manufacturing chains that in-

[5]These are analogous to the effects of public good production on tax revenue in a second-best world that are discussed by Atkinson and Stern (1974).

[6]See, *e.g.,* Atkinson and Sitglitz (1976, 1981).

[7]See equations (19), (20), and (22) of MP.

[8]See for instance, Barthold (1994); Metcalf (2009a, 2009b); the Mirrlees Review (2010, 2011) and the chapters therein, especially those by Fullerton, Leicester, and Smith (2010) and Crawford, Keen, and Smith (2010); Williams (2015); Levell, O'Connell, and Smith (2016).

clude the externality-causing goods in lieu of the externality-causing goods. This suggests that, for the purpose of externality control, taxes on goods forming a manufacturing chain are linked and cannot be chosen completely independently of one another. This aspect of externality-taxation that has important bearing on design of real-life policies has not been incorporated in existing theoretical works on optimal commodity taxation. For instance, in the numerical application of their theoretical analysis, Bovenberg and Goulder (1996) extend their theoretical model to include thirteen industries producing intermediate energy or material inputs. Carbon taxation of intermediate energy inputs of coal, crude oil, and natural gas is considered, while ignoring the fact that these goods are linked in manufacturing to outputs of other industries that are also studied by them and which could also be alternatively taxed to control externality generation. For example, coal is usually an input into industries such as electrical utilities, construction, metal and manufacturing, and the output of electrical utilities is employed to run electronic appliances used by consumers and firms; while crude oil is an input for the petroleum refineries, whose output in turn is employed for driving motor vehicles to meet freight and other transportation needs.

In this work, we develop a modelling and an analytical framework for a general equilibrium study of some principles of second-best externality taxation, which address several issues discussed above. The stylised model distinguishes between intermediate input and consumption taxation and employs a definition of a tax equilibrium that is also Nash in the externality dimensions. This model is employed to show that second-best optimal intermediate input taxation involves (i) direct taxation of the sale of externality-causing goods; or (ii) taxation of non-substitutable inputs or outputs produced by them in manufacturing chains that include the externality-generating goods; or (iii) taxation of inputs of some services which are complementary to the input usage of goods that generate externalities; or (iv) appropriate combinations of (i) to (iii).[9] Thus, this work provides a theoretical foundation for the intuitive arguments made in policy-oriented works about upstream and downstream taxation of externality.

As in MP, it is shown that second-best optimal consumption taxes involve both the conventional "many person Ramsey rule (MPRR)" component and an externality-correction component. However, unlike in MP, the former component captures only the balance between equity and efficiency objectives of the government and is independent of the latter component. The externality components of consumption and optimal intermediate input taxes are linked, indicating several degrees of freedom in choosing them at a second-best. The MPRR component of a commodity tax can be implemented as a VAT or GST, while its externality component can be implemented as an excise duty. It is intuitive that, unlike in the case of the VAT, firms should not be rebated on excises paid on their purchases of externality-causing inputs such as motoring fuel, coal, or ozone depleting chemicals, for then it would beat the very purpose of levying such excises, which is to control the generation of harmful external effects. We show that the tax credit to a producer who pays retail prices (*i.e.,* prices paid by consumers) when

[9] We argue in Section 2.1 that non-substitutability of some inputs in production is justified by the fundamental laws of thermodynamics such as the material balance conditions and the entropy law.

buying inputs should be equal to his GST payment (the MPRR component of the retail price) plus the excess of the externality excise paid under the consumption tax over the corresponding second-best intermediate input tax. The consumption VAT on vehicular services and services of electrical and electronic devices based on the MPRR principle, automatically ensures that the demands for electricity, motoring fuel, and road-services (which are derived from the demands for these former services) are also appropriately restricted in keeping with the equity and efficiency objectives of the government. Thus, for commodities such as electricity, motoring fuel, and road-services, the MPRR component is shown to be zero, leading to the prescription that GST on these commodities should be zero, and that any taxation of these goods is justified purely on grounds of externality correction.

Section 2 provides a detailed description of the model employed in this paper. Three types of externalities, *viz.,* consumption, congestion, and a fossil-fuel (carbon) externalities, are initially modelled and the concepts of effective producer and consumer prices are introduced based on the non-substitutability of some inputs in production and complementarities between some goods in consumption and production. Section 3 characterises and discusses the structure of second-best optimal intermediate input taxes, while Section 4 does the same for optimal consumption taxes. These results are stated in the forms of Theorems 6 and 8, which clearly bring out the degrees of freedom in choosing commodity tax rates given the linkages between various goods in manufacturing and consumption in our model. Detailed discussions of these results, their implications, and some comparisons with real-life commodity-tax policies are provided in Sections 3.2 and 4.2.[10] Exploiting the Nash elements in the definition of a tax equilibrium, Section 5 derives expressions for the social marginal costs of different externalities. Section 6 discusses the implementation of the optimal intermediate input and consumption taxes as a system of (consumption-oriented) VAT with input-tax credits. It derives the second-best optimal principles for rebating the tax paid by producers while purchasing intermediate-inputs. The model can be extended along several dimensions. In Section 7, it is extended to distinguish between two types of subsidy policies (a) Pigouvian subsidies that can be given to encourage generation of positive externalities emanating from services such as carbon capture and storage and (b) subsidies given to meet quantity targets for activities such as renewable energy generation and use of cleaner motoring fuels, which are set by the government to promote long-run goals such as sustainable development. We conclude in Section 8.

2 The model

2.1 Description of the goods and externalities in the economy.

We begin with a model containing three types of externalities – a consumption externality, a congestion externality, and a fossil-fuel/carbon externality, quantities of which are denoted by

[10] The reader can refer to Theorems 6 and 8 and read Sections 3.2 and 4.2 in isolation from the more technical Sections 3.1 and 4.1.

non-negative scalars z_X, z_C, and z_Y, respectively. There are N basic commodities available for final consumption. The index set of these goods is denoted by $\mathcal{N}$. These goods can also be used as intermediate inputs by firms.

Industrial production is governed by fundamental laws of thermodynamics, which dictate the rules by which material inputs are transformed into outputs.[11] These rules take the form of chemical equations and other engineering relations that satisfy the material-balance conditions and the entropy law. It follows from these rules that usage of certain goods as inputs are not only essential in the production of outputs of certain other goods, but also that the use of these inputs is highly correlated with the production of these outputs, *i.e.,* increases in production of these outputs necessarily require increased usage of these inputs.[12] For example, the production of iron (Fe) is highly correlated to the use of iron-ore (Fe_2O_3) as an input (as the chemical equation governing this transformation implies that two moles of the ore are necessarily required to produce four moles of iron.)[13] Similarly, the production of thermal electricity is highly correlated to the use of fossil-fuels such as coal, while the production of cigarettes (respectively, wine) is highly correlated to the use of tobacco (respectively, grapes). The basic laws of thermodynamic hence suggest that some goods are non-substitutable inputs in the production of outputs of some goods. Using the language of production theory in economics, we say that an input is *non-substitutable* in the production of output of a certain good if, starting from any technologically efficient production point, increase in the output of the good necessarily implies an increase in the usage of that input

We highlight below the goods that are associated with generation of the three externalities mentioned above:

- The consumption externality is caused by a good indexed by $\mathbb{l} \in \mathcal{N}$, representing a commodity such as cigarettes or alcohol. There is a non-substitutable input, indexed by $\mathbb{t} \in \mathcal{N}$, which is used in the production of good $\mathbb{l}$. For example, tobacco is a non-substitutable input in the production of cigarettes, while grapes is a non-substitutable input in the production of wine.

- A congestion externality (such as reduction in motoring/driving speed) is caused when services of motor vehicles are employed by consumers and firms in the presence of limited availability of road services at any point in time. The intensity to which congestion is caused varies with the size of the vehicles used. Thus, we distinguish between small, medium, and big-sized vehicles, which are indexed by $\mathbb{s}, \mathbb{m}, \mathbb{b} \in \mathcal{N}$, respectively. Road services are indexed by $\mathbb{R} \in \mathcal{N}$.

[11] See Baumgärtner and Arons (2003) for a very helpful discussion on the relationship between the fundamental laws of thermodynamics and modern industrial production.

[12] In standard production theory, an input is essential for the production of a good if no output of the good is produced when zero amount of that input is used.

[13] The chemical equation governing the transformation of iron ore into iron, where the energy for the reduction is provided by carbon in coal is

$$2\ Fe_2O_3\ +\ 3\ C \longrightarrow\ 4\ Fe\ +\ 3\ CO_2$$

- A fossil-fuel (carbon) externality is caused when coal or oil are combusted to meet consumption or production needs. Rendering of services by motor vehicles requires fuel as input. We index such a fuel (say, oil) by $\mathbb{o} \in \mathcal{N}$. While oil can be employed to run motor vehicles, it can also have non-motoring uses, which we index by $\mathbb{k} \in \mathcal{N}$. Coal available for commercial purposes and sold to both consumers and firms is indexed by $\mathbb{c} \in \mathcal{N}$. Commercial coal is a non-substitutable input in the production of non-renewable (thermal) electricity (which is indexed by $\mathbb{n}$) and in the production of a w-dimensional vector of goods such as iron and steel, bricks, cement, pulp and paper (which is indexed by $\mathbb{w} = \langle \mathbb{w}_1, \ldots, \mathbb{w}_w \rangle$, where $\mathbb{w}_i \in \mathcal{N}$ for all $i = 1, \ldots, w$).

- The production of refined oil $\mathbb{o}$ requires unrefined oil, indexed by $\mathbb{u}$, as a non-substitutable input. Coal is extracted from nature by mining. Coal extracted from the mines is treated as a non-substitutable input (indexed by $\mathbb{v}$) into the production of commercial coal $\mathbb{c}$ that is sold to consumers and firms. We assume that the goods $\mathbb{u}$ and $\mathbb{v}$ are inputs and not final goods that are consumed by consumers. Hence, $\mathbb{u}$ and $\mathbb{v}$ are not in the index set $\mathcal{N}$.[14]

- The model includes production of renewable electricity, which is indexed by $\mathbb{r}$, as an alternative to thermal electricity. The users of commercially produced electricity, *viz.,* firms and consumers, cannot distinguish between its renewable and non-renewable components. The homogeneous good electricity purchased by them is indexed by $\mathbb{e} \in \mathcal{N}$.

In what follows, it will often be convenient to use two types of partitions/splits of the index set $\mathcal{N}$:
(i) $\mathcal{N} \equiv \widetilde{\mathcal{N}} \cup \{\mathbb{w}_1, \ldots, \mathbb{w}_w, \mathbb{o}, \mathbb{l}, \mathbb{t}, \mathbb{c}, \mathbb{e}\}$, where set $\widetilde{\mathcal{N}}$ contains indices of $(N - 5 - w)$ goods in $\mathcal{N}$ other than the $w + 5$ goods $\mathbb{w}_1, \ldots, \mathbb{w}_w$, $\mathbb{o}$, $\mathbb{l}$, $\mathbb{t}$, $\mathbb{c}$, and $\mathbb{e}$
(ii) $\mathcal{N} \equiv \widehat{\mathcal{N}} \cup \{\mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}, \mathbb{R}, \mathbb{w}_1, \ldots, \mathbb{w}_w, \mathbb{o}, \mathbb{l}, \mathbb{t}, \mathbb{c}, \mathbb{e}\}$, where set $\widehat{\mathcal{N}}$ contains indices of all $(N - 10 - w)$ goods in $\mathcal{N}$ other than the $w + 10$ goods in the index set $\{\mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}, \mathbb{R}, \mathbb{w}_1, \ldots, \mathbb{w}_w, \mathbb{o}, \mathbb{l}, \mathbb{t}, \mathbb{c}, \mathbb{e}\}$.

2.2 Modelling production.

There are F firms indexed by f. To allow for intermediate input taxation, for every good $i \in \mathcal{N}$, we distinguish between its usage as an input by a firm and its production as an output by the firm. An output-input production vector of firm f is defined as $y^f = \langle y^f_O,\ y^f_I \rangle \ \in\ \mathbf{R}^{\mathfrak{N}}_+$, where $y^f_O \in \mathbf{R}^{\mathfrak{N}_O}_+$, $y^f_I \in \mathbf{R}^{\mathfrak{N}_I}_+$, and $\mathfrak{N} = \mathfrak{N}_O + \mathfrak{N}_I$. The vector of amounts of different outputs produced by firm f is given by the $\mathfrak{N}_O = N + 3$-dimensional vector[15]

$$y^f_O = \left\langle y^f_{O\widetilde{\mathcal{N}}},\ y^f_{O\mathbb{w}},\ y^f_{O\mathbb{t}},\ y^f_{O\mathbb{r}},\ y^f_{O\mathbb{n}},\ y^f_{O\mathbb{l}},\ y^f_{O\mathbb{c}},\ y^f_{O\mathbb{o}},\ y^f_{O\mathbb{v}},\ y^f_{O\mathbb{u}} \right\rangle \ \in\ \mathbf{R}^{\mathfrak{N}_O}_+.$$

[14] As will be seen later, this distinction between refined and unrefined oil, on the one hand, and mined and commercial coal, on the other, permits a bigger choice of upstream and downstream taxation of carbon to contain the externality it causes. See for instance Metcalf (2009).

[15] We distinguish between vector and matrix notation. The vector notation of a point $a \in \mathbf{R}^n$ is $a = \langle a_1, a_2, \ldots, a_n \rangle$, while the corresponding matrix notation is $a^\top = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$, where $\top$ denotes the transpose operation.

For example, $y^f_{O\mathbb{c}}$ denotes the output of commercial coal produced by firm f, while $y^f_{O\mathbb{v}}$ denotes the output of coal excavated by firm f.[16] If firm f does not produce a particular good, then its output of that good is always zero.

The usage of different inputs by firm f is given by the $\mathfrak{N}_I = N+3+w$-dimensional vector[17]

$$y^f_I = \left\langle y^f_{I\widehat{\mathbb{N}}},\ y^f_{I\mathbb{s}},\ y^f_{I\mathbb{m}},\ y^f_{I\mathbb{b}},\ y^f_{I\mathbb{k}},\ y^f_{I\mathbb{R}},\ y^f_{I\mathbb{w}},\ y^f_{I\mathbb{tl}},\ y^f_{I\mathbb{to}},\ y^f_{I\mathbb{e}},\ y^f_{I\mathbb{l}},\ y^f_{I\mathbb{cn}},\ y^f_{I\mathbb{cw}},\ y^f_{I\mathbb{o}},\ y^f_{I\mathbb{v}},\ y^f_{I\mathbb{u}} \right\rangle \in \mathbf{R}^{\mathfrak{N}_I}_+.$$

For example, $y^f_{I\mathbb{e}}$ is the input of electricity used by firm f during its production. If firm f processes mined coal to produce commercial coal then $y^f_{I\mathbb{v}}$ denotes its usage of raw coal. We allow for the possibility that good $\mathbb{t}$ can be employed both to produce the externality producing good $\mathbb{l}$ (this input usage of good $\mathbb{t}$ is indexed by $\mathbb{tl}$) and other non-externality producing goods (this input usage of good $\mathbb{t}$ is indexed by $\mathbb{to}$).[18] We also distinguish between use of coal in the non-renewable electricity generating sector (this usage will be indexed by $\mathbb{cn}$) and its other uses by firms to produce goods $\mathbb{w}_1, \ldots, \mathbb{w}_w$ (which will be indexed by the vector $\mathbb{cw} = \langle \mathbb{cw}_1, \ldots, \mathbb{cw}_w \rangle$). Input usage of a good is zero for any firm f if it does not use the good as an input.

The set of technological possibilities of firm f are contingent on the extent of fossil-fuel and congestion externalities it faces and is denoted by $Y^f(z_C, z_Y) \subset \mathbf{R}^{\mathfrak{N}}_+$. In Assumption 1 below, non-substitutability of some inputs in producing some outputs is captured by assuming that these inputs are required in fixed proportions to produce these outputs.[19] Thus, coal is assumed to be required in a fixed proportion $\phi_{\mathbb{cn}}$ to produce non-renewable electricity $\mathbb{n}$. For $i = 1, \ldots, w$, $\phi_{\mathbb{cw}_i}$ is the amount of commercial coal needed to produce one unit of good $\mathbb{w}_i$. Good $\mathbb{t}$ is required in a fixed proportion $\phi_{\mathbb{t}}$ to produce the consumption externality-causing good $\mathbb{l}$. Similarly, $\phi_{\mathbb{u}}$ and $\phi_{\mathbb{v}}$ denote the amounts of unrefined oil and mined coal required to produce a unit each of refined oil and commercial coal, respectively. For $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$, the amount of oil demanded per unit distance covered by vehicle type i is denoted by α_i, while $\alpha_{\mathbb{k}}$ denotes the average fuel-intensity of the aggregate non-motoring good $\mathbb{k}$ whose use as an input also requires oil. Similarly, for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$, the amount of road-service demanded per unit distance covered by vehicle type i is denoted by β_i.[20]

Assumption 1 *For all $f = 1, \ldots, F$ and for all $z_Y \geq 0$ and $z_C \geq 0$, the set $Y^f(z_C, z_Y) \subset \mathbf{R}^{\mathfrak{N}}_+$ is non-empty, closed, convex, satisfies input and output free disposability, and* [21]

[16] $y^f_{O\widetilde{\mathbb{N}}}$ denotes an output production vector whose elements are output quantities of all goods with indexes in set $\widetilde{\mathbb{N}}$.

[17] $y^f_{I\widehat{\mathbb{N}}}$ denotes input production vector whose elements are input quantities of all goods with indexes in set $\widehat{\mathbb{N}}$.

[18] For example, grapes can be be used to produce not only wine but also preserves such as jam or can be directly sold as a fruit to consumers.

[19] This approach to model input non-substitutability is motivated by the chemical and engineering relations based on material balance conditions and entropy laws that govern transformation of material inputs into outputs during industrial production.

[20] Congestion externality (such as reduction in driving speed) increases as demand for road-space increases, while fossil-fuel externality increases as demands for oil and commercial coal increase.

[21] The technology set $Y^f(z_C, z_Y)$ satisfies input and output free disposability if $\langle y^f_O, y^f_I \rangle \in Y^f(z_C, z_Y)$ and $\bar{y}^f_O \leq y^f_O$ and $\bar{y}^f_I \geq y^f_I$ then $\langle \bar{y}^f_O, \bar{y}^f_I \rangle \in Y^f(z_C, z_Y)$, *i.e.,* if a production plan is feasible under the technology then it also permits a production plan with no bigger amounts of outputs and no lesser amount of inputs.

$$
\begin{aligned}
& y^f = \left\langle y_O^f, y_I^f \right\rangle \in Y^f\left(z_C, z_Y\right) \implies \\
& y_{I\mathbb{cn}}^f \geq \phi_{\mathbb{cn}} y_{O\mathbb{n}}^f \;\wedge\; y_{I\mathbb{tl}}^f \geq \phi_{\mathbb{t}} y_{O\mathbb{l}}^f \quad\wedge\quad y_{I\mathbb{o}}^f \geq \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i y_{Ii}^f \quad\wedge\quad y_{I\mathbb{R}}^f \geq \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i y_{Ii}^f \\
\wedge \quad & y_{I\mathbb{u}}^f \geq \phi_{\mathbb{u}} y_{O\mathbb{c}}^f \quad\wedge\quad y_{I\mathbb{v}}^f \geq \phi_{\mathbb{v}} y_{O\mathbb{o}}^f \quad\wedge\quad y_{I\mathbb{cw}_i}^f \geq \phi_{\mathbb{w}_i} y_{O\mathbb{w}_i}^f \;\forall\; i = 1, \ldots, w.
\end{aligned}
$$

Thus, for example, if a production plan $y^f = \left\langle y_O^f, y_I^f \right\rangle$ is feasible under technology $Y^f\left(z_C, z_Y\right)$, then the amount of coal used as an input by this production plan for producing $y_{O\mathbb{n}}^f$ amount of non-renewable energy (which is given by $y_{I\mathbb{cn}}^f$) is at least as big as $\phi_{\mathbb{cn}} y_{O\mathbb{n}}^f$. The amount of oil used as an input under this production plan is at least as big as the sum of the amounts of oil required to run big, small, and medium-sized vehicles and the amount used for non-motoring purposes $\mathbb{k}$: $\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i y_{Ii}^f$.

By allowing us to distinguish between input and output uses of any good, the above model of a technology allows us to distinguish between the input and output price vectors faced by the producers, and hence permits us to define a vector of intermediate input taxes. The output-input producer price vector is denoted by $p = \langle p_O,\ p_I \rangle \in \mathbf{R}_+^{\mathfrak{N}}$, where the vector of output prices faced by firms is

$$
p_O = \left\langle p_{O\widetilde{\mathbb{N}}},\ p_{O\mathbb{w}},\ p_{O\mathbb{t}},\ p_{O\mathbb{r}},\ p_{O\mathbb{n}},\ p_{O\mathbb{l}},\ p_{O\mathbb{c}},\ p_{O\mathbb{o}},\ p_{O\mathbb{v}},\ p_{O\mathbb{u}} \right\rangle \in \mathbf{R}_{++}^{\mathfrak{N}_O}
$$

and the vector of input prices faced by firms is

$$
p_I = \left\langle p_{I\widehat{\mathbb{N}}},\ p_{I\mathbb{s}},\ p_{I\mathbb{m}},\ p_{I\mathbb{b}},\ p_{I\mathbb{k}},\ p_{I\mathbb{R}},\ p_{I\mathbb{w}},\ p_{I\mathbb{tl}},\ p_{I\mathbb{ta}},\ p_{I\mathbb{e}},\ p_{I\mathbb{l}},\ p_{I\mathbb{cn}},\ p_{I\mathbb{cw}},\ p_{I\mathbb{o}},\ p_{I\mathbb{v}},\ p_{I\mathbb{u}} \right\rangle \in \mathbf{R}_{++}^{\mathfrak{N}_I}.
$$

Thus, the vector of intermediate input tax rates associated with the producer price vector $p = \langle p_O,\ p_I \rangle \in \mathbf{R}_+^{\mathfrak{N}}$ is defined as the vector $\eta \in \mathbf{R}^{\mathfrak{N}_O + 2}$ with elements

$$
\begin{aligned}
\eta_i &= p_{Ii} - p_{Oi} && \forall\; i \in \widetilde{\mathbb{N}} \cup \{\mathbb{w}_1, \ldots, \mathbb{w}_w, \mathbb{l}, \mathbb{o}, \mathbb{v}, \mathbb{u}\} \\
\eta_{\mathbb{t}i} &= p_{I\mathbb{t}i} - p_{O\mathbb{t}} && \forall\; i = \mathbb{a}, \mathbb{l} \\
\eta_{\mathbb{cw}_i} &= p_{I\mathbb{cw}_i} - p_{O\mathbb{c}} && \forall\; i = 1, \ldots, w \\
\eta_{\mathbb{cn}} &= p_{I\mathbb{cn}} - p_{O\mathbb{c}} \\
\eta_i &= p_{I\mathbb{e}} - p_{Oi} && \forall\; i = \mathbb{r}, \mathbb{n}
\end{aligned} \tag{1}
$$

This is the vector of taxation of transactions between firms in commodities. In particular, based on the input usages of goods $\mathbb{t}$ and $\mathbb{c}$, transactions between firms in these goods can be differentially taxed. Thus, $\eta_{\mathbb{tl}}$ is the tax on transactions between firms producing good $\mathbb{t}$ and firms using this good to produce the externality-causing good $\mathbb{l}$, while $\eta_{\mathbb{ta}}$ is the tax on transactions between firms producing good $\mathbb{t}$ and firms using this good to produce non-externality causing goods. It also permits different intermediate-input taxation of renewable or non-renewable electricity ($\eta_{\mathbb{r}}$ versus $\eta_{\mathbb{n}}$). In this case, all firms purchasing electricity as an input face a common input price $p_{I\mathbb{e}}$. However, prices received by electricity-generating firms vary (as $p_{O\mathbb{r}}$ or $p_{O\mathbb{n}}$) depending on whether these firms are selling renewable or non-renewable electricity.

The externality-constrained profit maximisation problem of firm f is:

$$\max_{y^f \in \mathbf{R}^{\mathfrak{N}}{}_{+}} \quad \left\{ p_O y_O^f - p_I y_I^f \;\middle|\; y^f = \left\langle y_O^f, y_I^f \right\rangle \in Y^f\left(z_C, z_Y\right) \right\} \tag{2}$$

The problem can more usefully be redefined in terms of the following vectors of *effective* producer prices which are relevant when some inputs are non-substitutable in production of some outputs.

$$\begin{aligned}
\rho_O &= \left\langle \rho_{O\widetilde{\mathbb{N}}},\ \rho_{O\mathbb{w}_1},\ \ldots,\ \rho_{O\mathbb{w}_w},\ \rho_{O\mathbb{t}},\ \rho_{O\mathbb{r}},\ \rho_{O\mathbb{n}},\ \rho_{O\mathbb{l}},\ \rho_{O\mathbb{c}},\ \rho_{O\mathbb{o}},\ \rho_{O\mathbb{v}},\ \rho_{O\mathbb{u}} \right\rangle \in \mathbf{R}_{+}^{\mathfrak{N}_O} \equiv \mathbf{R}_{+}^{\mathfrak{N}_{\rho_O}} \\
&= \Big\langle p_{O\widetilde{\mathbb{N}}},\ p_{O\mathbb{w}_1} - \phi_{\mathbb{c}\mathbb{w}_1}\bar{p}_{I\mathbb{c}\mathbb{w}_1}, \ldots, p_{O\mathbb{w}_w} - \phi_{\mathbb{c}\mathbb{w}_w}\bar{p}_{I\mathbb{c}\mathbb{w}_w},\ p_{O\mathbb{t}},\ p_{O\mathbb{r}},\ \left(p_{O\mathbb{n}} - \phi_{\mathbb{c}\mathbb{n}} p_{I\mathbb{c}\mathbb{n}}\right),\ \left(p_{O\mathbb{l}} - \phi_{\mathbb{t}} p_{I\mathbb{t}\mathbb{l}}\right), \\
&\qquad \left(p_{O\mathbb{c}} - \phi_{\mathbb{v}} p_{I\mathbb{v}}\right),\ \left(p_{O\mathbb{o}} - \phi_{\mathbb{u}} p_{I\mathbb{u}}\right),\ p_{O\mathbb{v}},\ p_{O\mathbb{u}} \Big\rangle \\
\rho_I &= \left\langle \rho_{I\widehat{\mathbb{N}}},\ \rho_{I\mathbb{s}},\ \rho_{I\mathbb{m}},\ \rho_{I\mathbb{b}},\ \rho_{I\mathbb{k}},\ \rho_{I\mathbb{w}_1},\ \ldots,\ \rho_{I\mathbb{w}_w},\ \rho_{I\mathbb{t}\mathbb{a}},\ \rho_{I\mathbb{e}},\ \rho_{I\mathbb{l}} \right\rangle \in_{+}^{\mathfrak{N}_{\rho_I}} \equiv \mathbf{R}_{+}^{N-3} \\
&= \Big\langle p_{I\widehat{\mathbb{N}}},\ \left(p_{I\mathbb{s}} + \alpha_{\mathbb{s}} p_{I\mathbb{o}} + \beta_{\mathbb{s}} p_{I\mathbb{R}}\right),\ \left(p_{I\mathbb{m}} + \alpha_{\mathbb{m}} p_{I\mathbb{o}} + \beta_{\mathbb{m}} p_{I\mathbb{R}}\right),\ \left(p_{I\mathbb{b}} + \alpha_{\mathbb{b}} p_{I\mathbb{o}} + \beta_{\mathbb{b}} p_{I\mathbb{R}}\right), \\
&\qquad \left(p_{I\mathbb{k}} + \alpha_{\mathbb{k}} p_{I\mathbb{o}}\right),\ p_{I\mathbb{w}_1},\ \ldots,\ p_{I\mathbb{w}_w},\ p_{I\mathbb{t}\mathbb{a}},\ p_{I\mathbb{e}},\ p_{I\mathbb{l}} \Big\rangle \\
\rho &= \left\langle \rho_O, \rho_I \right\rangle \in \mathbf{R}_{+}^{\mathfrak{N}_{\rho_O}} + \mathbf{R}_{+}^{\mathfrak{N}_{\rho_I}} = \mathbf{R}_{+}^{\mathfrak{N}_{\rho}} \equiv \mathbf{R}_{++}^{2N}
\end{aligned} \tag{3}$$

Thus, the effective (net) producer price of an output that requires a non-substitutable input is defined as its market price minus the cost of the amount of the non-substitutable input needed to produce a unit of the output. For example, the effective price of a pack of cigarettes which requires $\phi_{\mathbb{t}}$ amount of tobacco is $\rho_{O\mathbb{l}} = p_{O\mathbb{l}} - \phi_{\mathbb{t}} p_{I\mathbb{t}\mathbb{l}}$. The effective producer price of an input is the sum of its market price and the costs of the amounts of all other inputs that are complementary to it. For example, to derive the input services of a vehicle for transportation, complementary inputs of oil and road services are required. Hence, the effective price faced by a firm for using the input services from a small-sized vehicle to transport its goods is: $\rho_{I\mathbb{s}} = p_{I\mathbb{s}} + \alpha_{\mathbb{s}} p_{I\mathbb{o}} + \beta_{\mathbb{s}} p_{I\mathbb{R}}$.

Given the definition of an effective price vector, we can redefine the externality-constrained profit function of every firm f as a mapping $\pi^f : \mathbf{R}_{+}^{\mathfrak{N}_{\rho}} \times \mathbf{R}_{+}^{2} \longrightarrow \mathbf{R}_{+}$ with image[22]

[22] Assumption 1 implies that, while defining the profit maximisation problem using the effective price vector, we can combine several terms in the objective function $p_O y_O^f - p_I y_I^f$, *e.g.*, $p_{O\mathbb{l}} y_{O\mathbb{l}}^f - p_{I\mathbb{t}\mathbb{l}} y_{I\mathbb{t}\mathbb{l}}^f$ can be rewritten as $\left(p_{O\mathbb{l}} - \phi_{\mathbb{t}} p_{O\mathbb{t}\mathbb{l}}\right) y_{O\mathbb{l}}^f = \rho_{O\mathbb{l}} y_{O\mathbb{l}}^f$, once we note from Assumption 1 that $y_{I\mathbb{t}\mathbb{l}}^f = \phi_{\mathbb{t}} y_{O\mathbb{l}}^f$.

$$
\begin{aligned}
&\pi^f\left(\rho_O, \rho_I, z_C, z_Y\right) := \\
&\max_{y^f \in \mathbf{R}^{\mathfrak{N}}{}_+} \quad \rho_{O\widetilde{\mathbb{N}}} y^f_{O\widetilde{\mathbb{N}}} + \sum_{i=1}^{w} \rho_{O\mathbb{w}_i} y^f_{O\mathbb{w}_i} + \rho_{O\mathbb{t}} y^f_{O\mathbb{t}} + \rho_{O\mathbb{c}} y^f_{O\mathbb{c}} + \rho_{O\mathbb{o}} y^f_{O\mathbb{o}} + \rho_{O\mathbb{r}} y^f_{O\mathbb{r}} + \rho_{O\mathbb{n}} y^f_{O\mathbb{n}} + \rho_{O\mathbb{l}} y^f_{O\mathbb{l}} \\
&\quad + \rho_{O\mathbb{v}} y^f_{O\mathbb{v}} + \rho_{O\mathbb{u}} y^f_{O\mathbb{u}} - \rho_{I\widehat{\mathbb{N}}} y^f_{I\widehat{\mathbb{N}}} - \rho_{I\mathbb{ta}} y_{I\mathbb{ta}} - \rho_{I\mathbb{e}} y^f_{I\mathbb{e}} - \rho_{I\mathbb{l}} y^f_{I\mathbb{l}} \sum_{i=1}^{w} \rho_{I\mathbb{w}_i} y^f_{I\mathbb{w}_i} - \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \rho_{Ii} y^f_{Ii} \\
&\qquad \text{subject to} \\
&\qquad y^f = \langle y^f_O, y^f_I \rangle \in Y^f\left(z_C, z_Y\right), \quad \text{where}
\end{aligned}
$$

$$
\begin{aligned}
&y^f_{I\mathbb{R}} = \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i y^f_{Ii}; \quad y^f_{I\mathbb{tl}} = \phi_{\mathbb{t}} y^f_{O\mathbb{l}}; \quad y^f_{I\mathbb{cw}_i} = \phi_{\mathbb{cw}_i} y^f_{O\mathbb{w}_i} \ \forall\, i = 1, \ \ldots, \ w; \quad y^f_{I\mathbb{cn}} = \ \phi_{\mathbb{cn}} y^f_{O\mathbb{n}}; \\
&y^f_{I\mathbb{o}} = \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i y^f_{Ii}; \quad y^f_{I\mathbb{v}} = \phi_{\mathbb{v}} y^f_{O\mathbb{c}}; \qquad y^f_{I\mathbb{u}} = \phi_{\mathbb{u}} y^f_{O\mathbb{o}},
\end{aligned} \tag{4}
$$

In what follows, we will be characterising the effective producer prices at a tax equilibrium and at a social optimum. In this regards, we argue below that (3) implies that several producer price vectors $p = \langle p_O, p_I \rangle \in \mathbf{R}^{\mathfrak{N}}_+$ are consistent with a given vector of effective producer prices, $\rho = \langle \rho_O, \rho_I \rangle \in \mathbf{R}^{\mathfrak{N}_\rho}_+$. Denote the set of all such producer price vectors as $\mathcal{P}\left(\rho_O, \rho_I\right) \subset \mathbf{R}^{\mathfrak{N}}_+$. It is clear that $p = \langle p_O, p_I \rangle \in \mathcal{P}\left(\rho_O, \rho_I\right)$ if and only if it solves

$$
\begin{aligned}
\left\langle p_{O\widetilde{\mathbb{N}}}, \ p_{O\mathbb{t}}, \ p_{O\mathbb{r}}, \ p_{O\mathbb{v}}, \ p_{O\mathbb{u}} \right\rangle &= \left\langle \rho_{O\widetilde{\mathbb{N}}}, \ \rho_{O\mathbb{t}}, \ \rho_{O\mathbb{r}}, \ \rho_{O\mathbb{v}}, \ \rho_{O\mathbb{u}} \right\rangle \\
\left\langle p_{I\widehat{\mathbb{N}}}, \ p_{I\mathbb{w}_1}, \ \ldots, \ p_{I\mathbb{w}_w}, \ p_{I\mathbb{ta}}, \ p_{I\mathbb{e}}, \ p_{I\mathbb{l}} \right\rangle &= \left\langle \rho_{I\widehat{\mathbb{N}}}, \ \rho_{I\mathbb{w}_1}, \ \ldots, \ \rho_{I\mathbb{w}_w}, \ \rho_{I\mathbb{ta}}, \ \rho_{I\mathbb{e}}, \ \rho_{I\mathbb{l}} \right\rangle \\
p_{O\mathbb{w}_i} - \phi_{\mathbb{cw}_i} p_{I\mathbb{cw}_i} &= \rho_{O\mathbb{w}_i} \quad \forall\, i = 1, \ldots, w, \qquad p_{O\mathbb{n}} - \phi_{\mathbb{cn}} p_{I\mathbb{cn}} \ = \ \rho_{O\mathbb{n}}, \\
p_{O\mathbb{l}} - \phi_{\mathbb{t}} p_{I\mathbb{tl}} &= \rho_{O\mathbb{l}}, \qquad p_{O\mathbb{c}} - \phi_{\mathbb{v}} p_{I\mathbb{v}} \ = \ \rho_{O\mathbb{c}}, \qquad p_{O\mathbb{o}} - \phi_{\mathbb{u}} p_{I\mathbb{u}} \ = \ \rho_{O\mathbb{o}}, \\
p_{Ii} + \alpha_i p_{I\mathbb{o}} + \beta_i p_{I\mathbb{R}} &= \rho_{Ii} \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \qquad p_{I\mathbb{k}} + \alpha_{\mathbb{k}} p_{I\mathbb{o}} \ = \ \rho_{I\mathbb{k}}
\end{aligned} \tag{5}
$$

It is clear that, given any effective producer price vector $\rho \in \mathbf{R}^{\mathfrak{N}_\rho}$, the input and output producer prices of some commodities are uniquely determined (*e.g.,* $p_{O\mathbb{r}}$ and $p_{I\mathbb{e}}$ are uniquely determined), while the output and input producer prices of other commodities are not unique (*e.g.,* $p_{O\mathbb{n}}$ and $p_{I\mathbb{cn}}$ are not uniquely determined. Any non-negative combination of the two that satisfies $p_{O\mathbb{n}} - \phi_{\mathbb{cn}} p_{I\mathbb{cn}} = \rho_{O\mathbb{n}}$ is valid). This also implies that the intermediate input tax vector associated with any effective producer price vector $\rho \in \mathbf{R}^{\mathfrak{N}_\rho}$ is not uniquely determined. Given any $p \in \mathcal{P}(\rho)$, the solution to problem (2) is the same as the solution to problem (4).

Assumption 2 *For all $f = 1, \ldots, F$, the profit function π^f is well-defined, takes positive values, and is twice continuously differentiable in the interior of its domain. Moreover, the fossil-fuel and congestion externalities are detrimental for the firm, i.e., $\frac{\partial \pi^f}{\partial z_i} < 0$ for $i = C, Y$.*[23]

[23] Note that if the fossil-fuel and congestion externalities are detrimental for firm f then the set of technologically feasible output-input production vectors shrinks when these externality levels increase: $Y^f\left(\bar{z}_C, \bar{z}_Y\right) \subset$

In addition, duality theory implies that the profit function π^f is convex and homogenous of degree one in ρ for all firms $f = 1, \ldots, F$. The index set of goods for which the effective producer prices are defined (*i.e.,* the elements of vector $\rho \in \mathbf{R}_+^{\mathfrak{N}_\rho}$) is given by[24]

$$\begin{aligned}\mathfrak{I}_\rho \;:=\; & \left\{O\widetilde{\mathcal{N}}, O\mathbb{w}_1, \ldots, O\mathbb{w}_w,\ O\mathbb{t}, O\mathbb{r}, O\mathbb{n}, O\mathbb{l}, O\mathbb{c}, O\mathbb{o}, O\mathbb{v}, O\mathbb{u}\right\} \bigcup \\ & \qquad \left\{I\widehat{\mathcal{N}}, I\mathbb{s}, I\mathbb{m}, I\mathbb{b}, I\mathbb{k}, I\mathbb{w}_1, \ldots, I\mathbb{w}_w,\ I\mathbb{ta}, I\mathbb{e}, I\mathbb{l}\right\}. \qquad (6)\end{aligned}$$

Employing the Hotelling's lemma the optimal output supply and input demand vectors of the firm for goods in index set $\mathfrak{I}_\rho$ can be expressed as:

$$\begin{aligned}\mathfrak{y}_O^f &= \left\langle y^f_{O\widetilde{\mathcal{N}}}, y^f_{O\mathbb{w}_1}, \ldots, y^f_{O\mathbb{w}_w}, y^f_{O\mathbb{t}}, y^f_{O\mathbb{r}}, y^f_{O\mathbb{l}}, y^f_{O\mathbb{c}}, y^f_{O\mathbb{o}}, y^f_{O\mathbb{v}}, y^f_{O\mathbb{u}}\right\rangle \\ &= \nabla_{\rho_O}\pi^f\left(\rho_O, \rho_I, z_C, z_Y\right) \in \mathbf{R}_+^{\mathfrak{N}_{\rho_O}} \\ \mathfrak{y}_I^f &= \left\langle -y^f_{I\widehat{\mathcal{N}}}, -y^f_{I\mathbb{s}}, -y^f_{I\mathbb{m}}, -y^f_{I\mathbb{b}}, -y^f_{I\mathbb{k}}, -y^f_{I\mathbb{w}_1}, \ldots, -y^f_{I\mathbb{w}_w}, -y^f_{I\mathbb{ta}}, -y^f_{I\mathbb{e}}, -y^f_{I\mathbb{l}}\right\rangle \\ &= \nabla_{\rho_I}\pi^f\left(\rho_O, \rho_I, z_C, z_Y\right) \in \mathbf{R}_+^{\mathfrak{N}_{\rho_I}} \\ \mathfrak{y}^f &= \left\langle \mathfrak{y}_O^f,\ \mathfrak{y}_I^f\right\rangle \in \mathbf{R}^{\mathfrak{N}_\rho}. \qquad (7)\end{aligned}$$

Given the above vector of input demands and output supplies $\mathfrak{y}^f$, the profit maximising demands of the remaining (*i.e.,* non-substitutable) inputs are derived from (4).

Following Murty (2013), we assume that the institutional rules of profit income distribution in the economy partition the F firms in the economy into G firm-groups $\mathcal{G}^1, \ldots, \mathcal{G}^G$ such that the profit income of any consumer is a function of the shares he holds in the profits of these G firm-groups, where the profit of any firm-group g is the sum of profits of all firms in this firm-group:

$$\Pi^g\left(\rho_O, \rho_I, z_C, z_Y\right) := \sum_{f \in G^g} \pi^f\left(\rho_O, \rho_I, z_C, z_Y\right)\ \forall\ g = 1, \ldots, G.$$

The share of the h^{th} consumer in the g^{th} firm-group is denoted by θ_g^h. We assume that the profit taxation power is perfectly aligned to this grouping, *i.e.,* the government can implement G profit tax rates, one for every firm-group. Murty (2013) showed that, in the absence of externalities, the above assumptions will imply that production efficiency is true at a second-best optimum. We explore whether this will continue to be true in the current model with externalities.

2.3 Modelling consumption.

There are H consumers indexed by h. The net-consumption vector of consumer h is denoted $x^h \in \mathbf{R}^N$. Consumer h's preferences are represented by a continuously differentiable and strictly quasi-concave utility function $u^h : \mathbf{R}^{N-1} \times \mathbf{R}_+^3 \longrightarrow \mathbf{R}$ with image:[25]

$Y^f\left(z_C, z_Y\right)$ whenever $\langle \bar{z}_C, \bar{z}_Y\rangle > \langle z_C, z_Y\rangle$.

[24] Note that the effective producer prices are not defined for all inputs. For example, they are not defined for input usage of $\mathbb{o}$ and $\mathbb{R}$ as costs of using these inputs are incorporated in the effective producer prices of services from various vehicles that use these inputs.

[25] $\mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\}$ stands for set comprising all elements in $\mathcal{N}$ excluding elements $\mathbb{o}$ and $\mathbb{R}$.

$$u^h = u^h\left(x^h_{\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}}, z_C, z_X, z_Y\right). \tag{8}$$

In the formulation of the utility function above, it is assumed that consumers derive utility from all goods in $\mathcal{N}$ except oil and road-service. In particular, consumers utility depends on services of vehicles and other non-motoring goods that require oil and/or road service to function. Hence, the consumption demands for oil and road-service are indirectly induced by consumer demands for vehicular services and services of other non-motoring goods. These are given by

$$x^h_{\mathbb{o}} \;=\; \alpha_{\mathbb{s}}x^h_{\mathbb{s}} + \alpha_{\mathbb{m}}x^h_{\mathbb{m}} + \alpha_{\mathbb{b}}x^h_{\mathbb{b}} + \alpha_{\mathbb{k}}x^h_{\mathbb{k}} \quad \text{and} \quad x^h_{\mathbb{R}} \;=\; \beta_{\mathbb{s}}x^h_{\mathbb{s}} + \beta_{\mathbb{m}}x^h_{\mathbb{m}} + \beta_{\mathbb{b}}x^h_{\mathbb{b}}. \tag{9}$$

The utility function shows that consumer preferences are also affected adversely by amounts of the three externalities in the model, namely, z_X, z_C, and z_Y.

Taking the level of all externalities as given and beyond his influence, each consumer h maximises his utility subject to the following budget constraint:

$$q_{\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}}x^h_{\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \;+\; q_{\mathbb{o}}\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}}\alpha_i x^h_i \;+\; q_{\mathbb{R}}\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}}\beta_i x^h_i \;\leq\; m + m^h, \quad \text{where}$$

$$m^h = \sum_g (1-\tau_g)\,\theta^h_g \Pi^g(\rho_O, \rho_I, z_C, z_Y),$$

$q \in \mathbf{R}^N_{++}$ is the vector of prices faced by the consumers and m is a uniform lump-sum transfer (also called a demogrant) to them. The profit tax rate implemented on the g^{th} firm-group is denoted by τ_g for $g = 1,\ldots,G$. Hence, m^h denotes the profit income of consumer h, who owns shares in different firm-groups. Employing (9), the above budget constraint can be reduced to the following effective budget constraint define in the space of commodities that directly enter consumer utility function, namely, $\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}$.

$$\sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}}\delta_i x^h_i \leq\; m + m^h$$

where

$$\delta_i = q_i \;\forall\; i \in \mathcal{N}\setminus\{\mathbb{o},\mathbb{R},\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}\}; \quad \delta_i = q_i + \alpha_i q_{\mathbb{o}} + \beta_i q_{\mathbb{R}} \;\forall\; i = \mathbb{s},\mathbb{m},\mathbb{b}; \quad \delta_{\mathbb{k}} = q_{\mathbb{k}} + \alpha_{\mathbb{k}} q_{\mathbb{o}} \tag{10}$$

Thus, $\delta = \langle \delta_{\mathcal{N}\setminus\{\mathbb{o},\mathbb{R},\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}\}}, \delta_{\mathbb{s}}, \delta_{\mathbb{m}}, \delta_{\mathbb{b}}, \delta_{\mathbb{k}}\rangle \in \mathbf{R}^{N-2}_+$ is the vector of effective consumer prices, and consumers choose amounts of goods in $\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}$ to maximise their respective utility subject to the above effective budget constraint. The resulting indirect utility function is $u^h = V^h\left(\delta, m + m^h, z_C, z_X, z_Y\right)$.

For all $h = 1,\ldots,H$, we define vectors $\chi^h := \left\langle x^h_{\widehat{\mathcal{N}}}, x^h_{\mathbb{s}}, x^h_{\mathbb{m}}, x^h_{\mathbb{b}}, x^h_{\mathbb{k}}, x^h_{\mathbb{w}}, x^h_{\mathbb{t}}, x^h_{\mathbb{e}}, x^h_{\mathbb{l}}, x^h_{\mathbb{c}}\right\rangle \in \mathbf{R}^{N-2}_+$ (*i.e.,* χ^h is the consumption vector x^h purged off the elements corresponding to oil and road-service). Assuming that λ^h denotes the marginal utility of money for consumer h, from the envelope theorem it follows that the net consumer demands of all final goods other than $\mathbb{o}$ and $\mathbb{R}$ can be expressed in terms of λ^h and the derivatives of the indirect utility function:

$$V^h_{\delta_i}\left(\delta, m + m^h, z_C, z_X, z_Y\right) = -\lambda^h \chi^h$$

Given the utility maximising demands for the above consumption goods, the demands for the remaining two goods, $\mathbb{o}$ and $\mathbb{R}$, can be derived using (9). Given a vector of effective consumer prices $\delta \in \mathbf{R}^{N-2}_{+}$, (10) shows that the consumer price vector associated with it is not unique. The set of consumer price vectors $q \in \mathbf{R}^N$ associated with a given vector of effective consumer prices $\delta \in \mathbf{R}^{N-2}_{+}$ is denoted by $\mathcal{Q}(\delta)$. It consists of all vectors $q \in \mathbf{R}^N_{+}$ that solve (10).

Given a consumer price vector $q \in \mathbf{R}^N_{++}$ and a producer price vector $p = \langle p_O, p_I \rangle \in \mathbf{R}^{\mathfrak{N}}_{++}$, we define the rates of total consumption taxes associated with q and p as the vector $t \in \mathbf{R}^N$ of differences between prices paid by the consumers and the prices received by producers when they sell their outputs:[26]

$$\begin{aligned} q_j &= p_{Oj} + t_j \quad \forall\, j \in \mathcal{N} \setminus \{\mathbb{e}\} \\ q_{\mathbb{e}} &= p_{I\mathbb{e}} + t_{\mathbb{e}} \end{aligned} \tag{11}$$

Since the consumer price vector corresponding to a vector of effective consumer prices is not uniquely determined, (11) shows that the vector of consumption taxes is also not uniquely determined for any given producer price vector $p \in \mathbf{R}^{\mathfrak{N}}_{++}$ and effective consumer price vector $\delta \in \mathbf{R}^{N-2}_{+}$.

2.4 A tax equilibrium.

In the context of the current model, a tax equilibrium is defined by a configuration of all effective prices in consumption and production and other government policy variables

$$\langle \delta, \rho_O, \rho_I, m, \tau_1, \ldots, \tau_G, z_C, z_X, z_Y \rangle \in \mathbf{R}^{N+\mathfrak{N}_\rho+G+2}$$

such that, for each commodity, aggregate demand is equal to the aggregate supply[27]

[26] Recall, there is no unique output price for electricity as our model permits non-renewable and renewable electricity producing firms to receive different prices. However, from the purchasers point of view, electricity is a homogeneous good whether produced from renewable or non-renewable sources; hence they pay a common price. Here we have adopted the rule that consumption tax on electricity is measured relative to its common input price $p_{I\mathbb{e}}$. This implies that, for electricity, the wedge between the consumer price and producer's output price is given by the sum of the consumption tax and the intermediate input tax: $q_i - p_{Oi} = [q_i - p_{I\mathbb{e}}] + [p_{I\mathbb{e}} - p_{Oi}] = t_{\mathbb{e}} + \eta_i$ for $i = \mathbb{r}, \mathbb{n}$. See also Figure 1.

[27] The definition of a tax equilibrium assumes that $m^h = \sum_g (1-\tau_g)\, \theta^h_g \Pi^g(\rho_O, \rho_I, z_Y)$ for all $h = 1, \ldots, H$. Recall also that we have assumed that unrefined oil and excavated coal are not directly demanded by consumers.

$$
\begin{aligned}
\sum_h x_i^h\left(\delta, m+m^h, z_C, z_X, z_Y\right) &= \sum_g y_{Oi}^g\left(\rho_O, \rho_I, z_C, z_Y\right) - \sum_g y_{Ii}^g\left(\rho_O, \rho_I, z_C, z_Y\right) \\
&\qquad\qquad \forall\ i \in \widetilde{\mathfrak{N}} \cup \{\mathbb{l}, \mathbb{o}, \mathbb{w}_1, \ldots, \mathbb{w}_w\} \\
\sum_h x_{\mathbb{t}}^h\left(\delta, m+m^h, z_C, z_X, z_Y\right) &= \sum_g y_{O\mathbb{t}}^g\left(\rho_O, \rho_I, z_C, z_Y\right) - \sum_g \sum_{i=\mathbb{a},\mathbb{l}} y_{I\mathbb{t}i}^g\left(\rho_O, \rho_I, z_C, z_Y\right) \\
\sum_h x_{\mathbb{c}}^h\left(\delta, m+m^h, z_C, z_X, z_Y\right) &= \sum_g y_{O\mathbb{c}}^g\left(\rho_O, \rho_I, z_C, z_Y\right) - \sum_g \sum_{i=\mathbb{n},\mathbb{w}_1,\ldots,\mathbb{w}_w} y_{I\mathbb{c}i}^g\left(\rho_O, \rho_I, z_C, z_Y\right) \\
\sum_h x_{\mathbb{e}}^h\left(\delta, m+m^h, z_C, z_X, z_Y\right) &= \sum_{i\in\{\mathbb{r},\mathbb{n}\}} \sum_g y_{Oi}^g\left(\rho_O, \rho_I, z_C, z_Y\right) - \sum_g y_{I\mathbb{e}}^g\left(\rho_O, \rho_I, z_C, z_Y\right) \\
0 &= \sum_g y_{Oi}^g\left(\rho_O, \rho_I, z_C, z_Y\right) - \sum_g y_{Ii}^g\left(\rho_O, \rho_I, z_C, z_Y\right) \quad \forall\ i = \mathbb{v}, \mathbb{u} \\
z_C = \mathcal{Z}_C\left(\rho_O, \rho_I, \delta, m+m^h, z_C, z_X, z_Y\right), &\qquad z_X = \mathcal{Z}_X\left(\delta, m+m^h, z_C, z_X, z_Y\right), \\
&\qquad\qquad z_Y = \mathcal{Z}_Y\left(\rho_O, \rho_I, \delta, m+m^h, z_C, z_X, z_Y\right). \qquad (12)
\end{aligned}
$$

The functions $\mathcal{Z}_C : \mathbf{R}_+^{\mathfrak{N}_\rho+N+2} \longrightarrow \mathbf{R}_+$, $\mathcal{Z}_X : \mathbf{R}_+^{N+2} \longrightarrow \mathbf{R}_+$, and $\mathcal{Z}_Y : \mathbf{R}_+^{\mathfrak{N}_\rho+N+2} \longrightarrow \mathbf{R}_+$, which are defined in the next section, give the levels of the consumption, congestion, and fossil-fuel externalities generated as a result of choices of various economic agents when they take the existing level of externalities, all prices (and hence all commodity tax rates), and values of other government policy variables as given. Thus, the last three equalities in (12) imply that that the tax equilibrium is a Nash equilibrium in the externality dimensions: The externalities generated when optimising agents choose their consumption and production plans taking the levels of externalities z_C, z_X and z_Y as given should also be equal to z_C, z_X and z_Y, respectively.

2.5 Modelling externality generation.

To keep the focus sharp on the issues we wish to explore in this paper, we simplify the analysis by assuming that externality generation in linearly related to the goods causing the externality.[28]

The congestion externality is attributed to the demand for road services by different vehicles for motoring purposes . It is given by the function[29]

$$
z_C = \mathcal{Z}_C\left(\rho_O, \rho_I, \delta, m+m^h, z_C, z_X, z_Y\right) = \psi_{\mathbb{R}} \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i \left[\sum_h x_i^h + \sum_g y_{Ii}^g\right], \qquad (13)
$$

where $\psi_{\mathbb{R}} > 0$ is the congestion-caused (for example, a decrease in driving time) due to an additional unit of demand for road-service. Thus, total congestion is $\psi_{\mathbb{R}}$ times total demand by

[28]Externality generation is also modelled linearly in many policy oriented works, with scientifically determined coefficients based on the laws of thermodynamics. For example, while measuring the external effects of chlorofluorocarbons (CFCs), different ozone-depleting factors are associated with different (CFCs) (see Barthold (1994)). IPCC provides CO_2 emission factors of various fossil-fuels.

[29]With an abuse of notation, in what follows, we will often ignore writing the arguments of the demands and supplies.

vehicles for road service.

Similarly, the consumption externality due to good $\mathbb{l}$ (*e.g.,* cigarettes or alcohol) is given by

$$z_X \;=\; \mathcal{Z}_X\left(\delta, m+m^h, z_C, z_X, z_Y\right) = \psi_{\mathbb{l}} \sum_h x^h_{\mathbb{l}}, \tag{14}$$

where $\psi_{\mathbb{l}} > 0$ is the amount of externality produced per unit of good $\mathbb{l}$ consumed.

The fossil-fuel externality is generated when coal and oil are consumed by households and used as inputs by firms for motoring, non-motoring, generation of electricity, and other purposes. The amount of the fossil-fuel externality is hence given by

$$\begin{aligned} z_Y \;&=\; \mathcal{Z}_Y\left(\rho_O, \rho_I, \delta, m+m^h, z_C, z_X, z_Y\right) \\ &=\; \psi_{\mathbb{o}} \left[\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i \left(\sum_h x^h_i + \sum_g y^g_{Ii} \right) \right] \;+\; \psi_{\mathbb{c}} \left[\sum_h x^h_{\mathbb{c}} + \sum_{i=1}^{w} \phi_{\mathbb{c}\mathbb{w}_i} \sum_g y^g_{O\mathbb{w}_i} + \phi_{\mathbb{c}\mathbb{n}} \sum_g y^g_{O\mathbb{n}} \right]. \end{aligned} \tag{15}$$

2.6 The welfare maximisation problem.

Let $W\left(u^1, \ldots, u^H\right)$ be an individualistic social welfare function, which is increasing in its arguments. The second-best welfare maximisation problem is the following:

$$\begin{gathered} \max_{\{\delta,\rho_O,\rho_I,m,\tau_1,\ldots,\tau_G,z_C,z_X,z_Y\}} \; W\left(V^1, \ldots, V^H\right) \\ \text{subject to} \\ \langle \delta,\ \rho_O, \rho_I,\ m,\ \tau_1, \ldots, \tau_G,\ z_C,\ z_X,\ z_Y \rangle \text{ satisfies (12)} \end{gathered} \tag{16}$$

The Lagrangian of the welfare maximisation problem is

$$\begin{aligned} L \;=\;& W\left(V^1, \ldots, V^H\right) - \nu^{\top}_{\widetilde{\mathbb{N}}} \left[\sum_h x^h_{\widetilde{\mathbb{N}}} - \sum_g y^g_{O\widetilde{\mathbb{N}}} + \sum_g y^g_{I\widetilde{\mathbb{N}}} \right] - \sum_{i=\mathbb{l},\mathbb{o},\mathbb{w}_1,\ldots,\mathbb{w}_w} \nu_i \left[\sum_h x^h_i - \sum_g y^g_{Oi} + \sum_g y^g_{Ii} \right] \\ & -\nu_{\mathbb{c}} \left[\sum_h x^h_{\mathbb{c}} - \sum_g y^g_{O\mathbb{c}} + \sum_{i\in\{\mathbb{n},\mathbb{w}_1,\ldots,\mathbb{w}_w\}} \sum_g y^g_{I\mathbb{c}i} \right] - \nu_{\mathbb{t}} \left[\sum_h x^h_{\mathbb{t}} - \sum_g y^g_{O\mathbb{t}} + \sum_{i\in\{\mathbb{l},\mathbb{a}\}} \sum_g y^g_{I\mathbb{t}i} \right] \\ & -\nu_{\mathbb{e}} \left[\sum_h x^h_{\mathbb{e}} - \sum_{i\in\{\mathbb{r},\mathbb{n}\}} \sum_g y^g_{Oi} - \sum_g y^g_{I\mathbb{e}} \right] - \sum_{i\in\{\mathbb{v},\mathbb{u}\}} \nu_i \left[-\sum_g y^g_{Oi} + \sum_g y^g_{I_1} \right] \\ & -\gamma_C \left[\mathcal{Z}_C\left(\rho_O, \rho_I, q, m+m^h, z_C, z_X, z_Y\right) - z_C \right] - \gamma_X \left[\mathcal{Z}_X\left(q, m+m^h, z_C, z_X, z_Y\right) - z_X \right] \\ & \quad -\gamma_Y \left[\mathcal{Z}_Y\left(\rho_O, \rho_I, q, m+m^h, z_C, z_X, z_Y\right) - z_Y \right]. \end{aligned}$$

where we define the vector of Lagrange multipliers for the non-externality constraints as $\nu = \left\langle \nu_{\widetilde{\mathbb{N}}}, \nu_{\mathbb{w}_1}, \ldots, \nu_{\mathbb{w}_w}, \nu_{\mathbb{t}}, \nu_{\mathbb{e}}, \nu_{\mathbb{l}}, \nu_{\mathbb{c}}, \nu_{\mathbb{o}}, \nu_{\mathbb{v}}, \nu_{\mathbb{u}} \right\rangle \in \mathbf{R}^{N+2}_{+}$, while the Lagrange multipliers for the externality constraints are denoted by scalars γ_X, γ_C, and γ_Y. In what follows, we will assume that the vector

$$\bar{\Upsilon} = \langle \bar{\delta},\ \bar{m},\ \bar{\rho}_O,\ \bar{\rho}_I,\ \bar{\tau}_1,\ \ldots,\ \bar{\tau}_G,\ \bar{z}_C,\ \bar{z}_X,\ \bar{z}_Y,\ \bar{\nu}_{\widetilde{\mathbb{N}}},\ \bar{\nu}_{\mathbb{w}},\ \bar{\nu}_{\mathbb{t}},\ \bar{\nu}_e,\ \bar{\nu}_{\mathbb{l}},\ \bar{\nu}_c,\ \bar{\nu}_{\mathbb{o}},\ \bar{\nu}_{\mathbb{v}},\ \bar{\nu}_{\mathbb{u}},\ \bar{\gamma}_C,\ \bar{\gamma}_X,\ \bar{\gamma}_Y \rangle$$

consists of a second-best optimal tax equilibrium configuration along with the values of the Lagrange multipliers of problem (16) evaluated at this optimum. We further assume that $\bar{\Upsilon}$ corresponds to an interior optimum. From (5), the set of producer price vectors corresponding to $\bar{\Upsilon}$ is $\mathcal{P}(\bar{\rho}_O, \bar{\rho}_I)$. Similarly, employing (10), the set of consumer price vectors corresponding to $\bar{\Upsilon}$ is $\mathcal{Q}(\bar{\delta})$.

As is standard in the optimal commodity tax literature, we define the social marginal utility of consumer h as $\mu^h := \frac{\partial W}{\partial V^h}$. Recalling that the private marginal utility of consumer h is $\lambda^h := \frac{\partial V^h}{\partial m}$, the social marginal utility of income to consumer h is given by $\mu^h \lambda^h$. Intuitively, this measures the increase in social welfare when a dollar of income is given to consumer h.

3 Optimal intermediate input tax rates.

We now study the optimal intermediate input tax rates or, equivalently, the optimal rates of taxation of commodity transactions between firms at the second-best $\bar{\Upsilon}$.

3.1 Characterisation of optimal intermediate input tax rates.

We first derive the first-order conditions (FOCs) of the second-best welfare maximisation problem with respect to effective producer prices ρ_i for $i \in \mathfrak{I}_\rho$ and the profit tax rates τ_g for $g = 1, \ldots . \mathcal{G}$.[30]

The FOC with respect to profit tax rate τ_g is

$$-\bar{\Pi}^g \sum_h A_g^h = 0, \tag{17}$$

where, for all $h = 1, \ldots, H$ and $g = 1, \ldots, \mathcal{G}$,

$$\begin{aligned} A_g^h \;=\; \theta_g^h \Big[& \mu_h \lambda_h - \bar{\nu}_{\tilde{\mathbb{N}}}^\top \nabla_{m^h} x_{\tilde{\mathbb{N}}}^h - \bar{\nu}_{\mathbb{w}} \nabla_{m^h} x_{\mathbb{w}}^h - \bar{\nu}_{\mathbb{l}} \nabla_{m^h} x_{\mathbb{l}}^h - \bar{\nu}_{\mathbb{t}} \nabla_{m^h} x_{\mathbb{t}}^h - \bar{\nu}_{\mathbb{c}} \nabla_{m^h} x_{\mathbb{c}}^h - \bar{\nu}_{\mathbb{e}} \nabla_{m^h} x_{\mathbb{e}}^h \\ & - \bar{\nu}_{\mathbb{R}} \nabla_{m^h} \bar{x}_{\mathbb{R}}^h - \bar{\gamma}_C \frac{\partial \mathcal{Z}_C}{\partial m^h} - \bar{\gamma}_X \frac{\partial \mathcal{Z}_X}{\partial m^h} - \bar{\gamma}_Y \frac{\partial \mathcal{Z}_Y}{\partial m^h} \Big] \end{aligned}$$

is the net social marginal benefit from an increase in income to consumer h due to a unit increase in net-of-tax profit of firm-group g. The FOC (17) says that, at a second-best optimum, the net social marginal benefit from changes in incomes of all consumers due to a unit increase in the profit tax rate applicable to firm-group g is zero.

Under Assumption 2, $\bar{\Pi}^g \neq 0$ for all firm-groups $g = 1, \ldots, \mathcal{G}$. Hence, (17) implies that $\sum_h A_g^h = 0$ for all $g = 1, \ldots, \mathcal{G}$. Employing this and recalling (4), it can be shown that the

[30] Recall the index set of commodities $\mathfrak{I}_\rho$ defined in (6). Note also that all FOCs are evaluated at the optimal solution configuration $\bar{\Upsilon}$ of the second-best welfare maximisation problem. Profit of firm-group g evaluated at $\bar{\Upsilon}$ is denoted by $\bar{\Pi}^g$ for all $g = 1, \ldots, \mathcal{G}$.

FOC with respect to ρ_i for $i \in \mathfrak{I}_\rho$ can be written as

$$\begin{aligned}
&\sum_{j\in\widetilde{\mathbb{N}}} \bar{\nu}_j \sum_g \frac{\partial y^g_{Oj}}{\partial \rho_i} + \sum_{j\in\{\mathbb{t},\mathbb{v},\mathbb{u}\}} \bar{\nu}_j \sum_g \frac{\partial y^g_{Oj}}{\partial \rho_i} + \bar{\nu}_{\mathbb{e}} \sum_{i\in\{\mathbb{r},\mathbb{n}\}} \sum_g \frac{\partial y^g_{Oi}}{\partial \rho_i} + \sum_{j=1}^{w} \left(\bar{\nu}_{\mathbb{w}_j} - \phi_{\mathbb{c}\mathbb{w}_j}\bar{\nu}_{\mathbb{c}}\right) \sum_g \frac{\partial y^g_{O\mathbb{w}_j}}{\partial \rho_i} \\
&\quad + (\bar{\nu}_{\mathbb{l}} - \phi_{\mathbb{t}}\bar{\nu}_{\mathbb{t}}) \sum_g \frac{\partial y^g_{O\mathbb{l}}}{\partial \rho_i} + (\bar{\nu}_{\mathbb{e}} - \phi_{\mathbb{c}\mathbb{n}}\bar{\nu}_{\mathbb{c}}) \sum_g \frac{\partial y^g_{O\mathbb{n}}}{\partial \rho_i} + (\bar{\nu}_{\mathbb{c}} - \phi_{\mathbb{v}}\bar{\nu}_{\mathbb{v}}) \sum_g \frac{\partial y^g_{O\mathbb{c}}}{\partial \rho_i} + (\bar{\nu}_{\mathbb{o}} - \phi_{\mathbb{u}}\bar{\nu}_{\mathbb{u}}) \sum_g \frac{\partial y^g_{O\mathbb{o}}}{\partial \rho_i} \\
&\quad - \sum_{j\in\widehat{\mathbb{N}}} \bar{\nu}_j \sum_g \frac{\partial y^g_{Ij}}{\partial \rho_i} - \sum_{j=1}^{w} \bar{\nu}_{\mathbb{w}_j} \sum_g \frac{\partial y^g_{I\mathbb{w}_j}}{\partial \rho_i} - \bar{\nu}_{\mathbb{t}} \sum_g \frac{\partial y^g_{I\mathbb{t}\mathbb{a}}}{\partial \rho_i} - \bar{\nu}_{\mathbb{e}} \sum_g \frac{\partial y^g_{I\mathbb{e}}}{\partial \rho_i} \\
&\qquad - \sum_{j=\mathbb{s},\mathbb{m},\mathbb{b}} (\bar{\nu}_j + \alpha_j \bar{\nu}_{\mathbb{o}} + \beta_j \bar{\nu}_{\mathbb{R}}) \frac{\partial y^g_{Ij}}{\partial \rho_i} - (\bar{\nu}_{\mathbb{k}} + \alpha_{\mathbb{k}}\bar{\nu}_{\mathbb{o}}) \frac{\partial y^g_{I\mathbb{k}}}{\partial \rho_i} - \bar{\gamma}_C \frac{\partial \mathcal{Z}_C}{\partial \rho_i} - \bar{\gamma}_Y \frac{\partial \mathcal{Z}_Y}{\partial \rho_i} = 0. \qquad (18)
\end{aligned}$$

But recalling (7), the FOCs in (18) can be re-written in matrix notation as[31]

$$\nabla^\top_\rho \mathfrak{y}\, \vartheta \;=\; \bar{\gamma}_C \nabla_\rho \mathcal{Z}_C + \bar{\gamma}_Y \nabla_\rho \mathcal{Z}_Y \qquad (19)$$

where, recalling the definition of $\mathfrak{y}$ in (7), we have [32]

$$\begin{aligned}
\nabla^\top_\rho \mathfrak{y} &= \begin{bmatrix} \nabla^\top_\rho \mathfrak{y}_O & \nabla^\top_\rho \mathfrak{y}_I \end{bmatrix} = \nabla^2_{\rho\rho}\Pi \\
\nabla^\top_\rho \mathfrak{y}_O &= \begin{bmatrix} \nabla^\top_\rho y_{O\widetilde{\mathbb{N}}} & \nabla_\rho y_{O\mathbb{w}} & \nabla_\rho y_{O\mathbb{t}} & \nabla_\rho y_{O\mathbb{r}} & \nabla_\rho y_{O\mathbb{n}} & \nabla_\rho y_{O\mathbb{l}} & \nabla_\rho y_{O\mathbb{c}} & \nabla_\rho y_{O\mathbb{o}} & \nabla_\rho y_{O\mathbb{v}} & \nabla_\rho y_{O\mathbb{u}} \end{bmatrix} \\
\nabla^\top_\rho \mathfrak{y}_I &= \begin{bmatrix} -\nabla^\top_\rho y_{I\widehat{\mathbb{N}}} & -\nabla_\rho y_{I\mathbb{s}} & -\nabla_\rho y_{I\mathbb{m}} & -\nabla_\rho y_{I\mathbb{b}} & -\nabla_\rho y_{I\mathbb{k}} & -\nabla_\rho y_{I\mathbb{w}} & -\nabla_\rho y_{I\mathbb{t}\mathbb{a}} & -\nabla_\rho y_{I\mathbb{e}} & -\nabla_\rho y_{I\mathbb{l}} \end{bmatrix} \\
\vartheta^\top &= \begin{bmatrix} \vartheta^\top_O & \vartheta^\top_I \end{bmatrix} \\
\vartheta^\top_O &= \begin{bmatrix} \bar{\nu}^\top_{\widetilde{\mathbb{N}}} & (\bar{\nu}_{\mathbb{w}} - \phi_{\mathbb{c}\mathbb{w}}\bar{\nu}_{\mathbb{c}}) & \bar{\nu}_{\mathbb{t}} & \bar{\nu}_{\mathbb{e}} & (\bar{\nu}_{\mathbb{e}} - \phi_{\mathbb{c}\mathbb{n}}\bar{\nu}_{\mathbb{c}}) & (\bar{\nu}_{\mathbb{l}} - \phi_{\mathbb{t}}\bar{\nu}_{\mathbb{t}}) & (\bar{\nu}_{\mathbb{c}} - \phi_{\mathbb{v}}\bar{\nu}_{\mathbb{v}}) & (\bar{\nu}_{\mathbb{o}} - \phi_{\mathbb{u}}\bar{\nu}_{\mathbb{u}}) & \bar{\nu}_{\mathbb{v}} & \bar{\nu}_{\mathbb{u}} \end{bmatrix} \\
\vartheta^\top_I &= \begin{bmatrix} \bar{\nu}^\top_{\widehat{\mathbb{N}}} & \bar{\nu}_{\mathbb{s}} + \alpha_{\mathbb{s}}\bar{\nu}_{\mathbb{o}} + \beta_{\mathbb{s}}\bar{\nu}_{\mathbb{R}} & \bar{\nu}_{\mathbb{m}} + \alpha_{\mathbb{m}}\bar{\nu}_{\mathbb{o}} + \beta_{\mathbb{m}}\bar{\nu}_{\mathbb{R}} & \bar{\nu}_{\mathbb{b}} + \alpha_{\mathbb{b}}\bar{\nu}_{\mathbb{o}} + \beta_{\mathbb{b}}\bar{\nu}_{\mathbb{R}} & \bar{\nu}_{\mathbb{k}} + \alpha_{\mathbb{k}}\bar{\nu}_{\mathbb{o}} & \bar{\nu}_{\mathbb{w}} & \bar{\nu}_{\mathbb{t}} & \bar{\nu}_{\mathbb{e}} & \bar{\nu}_{\mathbb{l}} \end{bmatrix} \\
\bar{\gamma}_C \nabla_\rho \mathcal{Z}_C &= \bar{\gamma}_C\ \psi_{\mathbb{R}} \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i \nabla_\rho y_{Ii} \\
\bar{\gamma}_Y \nabla_\rho \mathcal{Z}_Y &= \bar{\gamma}_Y\ \psi_{\mathbb{o}} \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i \nabla_\rho y_{Ii} \ +\ \bar{\gamma}_Y\ \psi_{\mathbb{c}} \left(\sum_{i=1}^{w} \phi_{\mathbb{c}\mathbb{w}_i} \nabla_\rho y_{O\mathbb{w}_i} \ +\ \phi_{\mathbb{c}\mathbb{n}} \nabla_\rho y_{O\mathbb{n}} \right) \qquad (20)
\end{aligned}$$

The first condition in (20) follows from the Hotelling's lemma in (7), while the last two conditions follow from (13) and (15).

Note that the equation system in (12) that defines a tax equilibrium is homogenous of degree zero in q, ρ, and m as net demands of consumers and the supplies and demands of producers are homogeneous of degree zero in these variables. Hence, the system permits one normalisation.

Assumption 3 *The vector of effective producer prices* $(\bar{\rho})$ *at the second-best optimum* $\bar{\Upsilon}$ *is normalised.*

[31] $\mathfrak{y} := \sum_{g=1}^{G} \mathfrak{y}^g \in \mathbf{R}_+^{\mathfrak{N}_{\rho O}} \times \mathbf{R}_-^{\mathfrak{N}_{\rho I}}$.

[32] $\Pi = \sum_{g=1}^{G} \Pi^g$. Note that $\nabla_\rho \mathcal{Z}_C$, $\nabla_\rho \mathcal{Z}_Y$ and $\nabla^\top_\rho \mathfrak{y}$ are evaluated at the second-best optimum $\bar{\Upsilon}$.

Since the vector of producer supplies and demands $\mathfrak{y} = \langle \mathfrak{y}_O,\ \mathfrak{y}_I \rangle \in \mathbf{R}^{\mathfrak{N}_\rho}$ is homogeneous of degree zero in ρ, the vector $\bar{\rho}$ lies in the null space of the square matrix $\nabla_\rho \mathfrak{y}$ evaluated at $\bar{\Upsilon}$.[33] Since $\nabla_\rho \mathfrak{y} = \nabla^2_{\rho\rho}\Pi$ is a symmetric matrix, $\bar{\rho}$ also lies in the null space of matrix $\nabla^\top_\rho \mathfrak{y}$. Hence, the matrix $\nabla^\top_\rho \mathfrak{y}$ is not a full-ranked matrix. The following is a common assumption in the public economics literature with respect to the Hessian of the profit function $\nabla^2_{\rho\rho}\Pi$.[34]

Assumption 4 *The rank of the $\mathfrak{N}_\rho$-dimensional square matrix $\nabla^\top_\rho \mathfrak{y}$ is $\mathfrak{N}_\rho - 1$.*[35]

The lemma below, whose proof is given in the appendix, is a consequence of a basic result in linear algebra.

Lemma 5 *If there exists $\bar{\sigma} \in \mathbf{R}^{\mathfrak{N}_\rho}$ that solves the linear equation system*

$$\nabla^\top_\rho \mathfrak{y}\ \sigma = \bar{\gamma}_C \nabla_\rho \mathfrak{Z}_C + \bar{\gamma}_Y \nabla_\rho \mathfrak{Z}_Y \tag{$*$}$$

and $\bar{\sigma} \neq \vartheta$, then there exists a non-zero scalar $\bar{\kappa}$ such that

$$\vartheta = \bar{\sigma} + \bar{\kappa}\bar{\rho}. \tag{21}$$

To apply this lemma to derive the optimal system of intermediate input tax rates of our model, we define $\bar{\sigma}$ as the following sum of vectors in $\mathbf{R}^{\mathfrak{N}_\rho}$:

$$\bar{\sigma} := \sum_{j=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \bar{\sigma}^{Yj} + \sum_{j=\mathbb{s},\mathbb{m},\mathbb{b}} \bar{\sigma}^{Cj} + \sum_{j=1}^{w} \bar{\sigma}^{\mathbb{w}_j} + \bar{\sigma}^{\mathbb{n}}, \qquad \text{where} \tag{22}$$

$$\begin{array}{rclllrcll}
\bar{\sigma}^{Yj}_i & = & 0 & \forall\, i \in \mathfrak{I}_\rho & i \neq j & \wedge & \bar{\sigma}^{Yj}_j & = & -\bar{\gamma}_Y \psi_{\mathbb{o}} \alpha_j \quad \forall\, j \in \{I\mathbb{s}, I\mathbb{m}, I\mathbb{b}, I\mathbb{k}\} \\
\bar{\sigma}^{Cj}_i & = & 0 & \forall\, i \in \mathfrak{I}_\rho & i \neq j & \wedge & \bar{\sigma}^{Cj}_j & = & -\bar{\gamma}_C \psi_{\mathbb{R}} \beta_j \quad \forall\, j \in \{I\mathbb{s}, I\mathbb{m}, I\mathbb{b}\} \\
\bar{\sigma}^{\mathbb{w}_j}_i & = & 0 & \forall\, i \in \mathfrak{I}_\rho & i \neq O\mathbb{w}_j & \wedge & \bar{\sigma}^{\mathbb{w}_j}_{O\mathbb{w}_j} & = & \bar{\gamma}_Y \psi_{\mathbb{c}} \phi_{\mathbb{c}\mathbb{w}_j} \quad \forall\, j = 1,\ldots,w \\
\bar{\sigma}^{\mathbb{n}}_i & = & 0 & \forall\, i \in \mathfrak{I}_\rho & i \neq O\mathbb{n} & \wedge & \bar{\sigma}^{\mathbb{n}}_{O\mathbb{n}} & = & \bar{\gamma}_Y \psi_{\mathbb{c}} \phi_{\mathbb{c}\mathbb{n}}
\end{array}$$

With this definition, $\bar{\sigma}$ solves the equation system $(*)$ in Lemma 5.[36] We can now apply the conclusion of Lemma 5. Hence, applying (21) and using the definitions of ϑ in (20) and $\bar{\sigma}$ in

[33] The null space of matrix $\nabla_\rho \mathfrak{y}$ is the subspace: $Null\,(\nabla_\rho \mathfrak{y}) := \left\{ \zeta \in \mathbf{R}^{\mathfrak{N}_\rho}_+ \;\middle|\; \nabla_\rho \mathfrak{y}\ \zeta = 0^{\mathfrak{N}_\rho} \right\}$.

[34] See for example, Guesnerie (1995) and Weymark (1979).

[35] By the fundamental theorem of linear algebra, the dimension of the null space of an n-dimensional square matrix is n minus its rank.

[36] This is because, the structures of $\nabla^\top_\rho \mathfrak{y}$, $\bar{\gamma}_C \nabla_\rho \mathfrak{Z}_C$, and $\bar{\gamma}_Y \nabla_\rho \mathfrak{Z}_Y$ (as seen in (20)) imply that

$$\begin{aligned}
\nabla^\top_\rho \mathfrak{y} \sum_{i=I\mathbb{s},I\mathbb{m},I\mathbb{b},I\mathbb{k}} \bar{\sigma}^{Yi} &= \bar{\gamma}_Y \psi_{\mathbb{o}} \sum_{j=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_j \nabla_\rho y_{Ij} \\
\nabla^\top_\rho \mathfrak{y} \sum_{i=I\mathbb{s},I\mathbb{m},I\mathbb{b}} \bar{\sigma}^{Ci} &= \bar{\gamma}_C \psi_{\mathbb{R}} \sum_{j=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_j \nabla_\rho y_{Ij} \\
\nabla^\top_\rho \mathfrak{y} \sum_{i=1}^{w} \bar{\sigma}^{\mathbb{c}\mathbb{w}_i} &= \bar{\gamma}_Y \psi_{\mathbb{c}} \sum_{i=1}^{w} \phi_{\mathbb{c}\mathbb{w}_i} \nabla_\rho y_{I\mathbb{c}\mathbb{w}_i} \\
\nabla^\top_\rho \mathfrak{y}\ \bar{\sigma}^{\mathbb{n}} &= \bar{\gamma}_Y \psi_{\mathbb{c}} \phi_{\mathbb{c}} \nabla_\rho y_{O\mathbb{n}}
\end{aligned}$$

(22), we have

$$
\begin{aligned}
\bar{\nu}_i &= \bar{\kappa}\bar{\rho}_{Oi} \;\forall\; i \in \widetilde{\mathbb{N}} \cup \{\mathbb{t}, \mathbb{v}, \mathbb{u}\}, \qquad \bar{\nu}_{\mathbb{e}} = \bar{\kappa}\bar{\rho}_{O\mathbb{r}}, \\
\bar{\nu}_{\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\bar{\nu}_{\mathbb{c}} &= \bar{\kappa}\bar{\rho}_{O\mathbb{w}_i} + \bar{\gamma}_Y \psi_{\mathbb{c}} \phi_{\mathbb{cw}_i} \;\forall\; i = 1, \ldots, w, \\
\bar{\nu}_{\mathbb{e}} - \phi_{\mathbb{cn}}\bar{\nu}_{\mathbb{c}} &= \bar{\kappa}\bar{\rho}_{O\mathbb{n}} + \bar{\gamma}_Y \psi_{\mathbb{c}} \phi_{\mathbb{cn}}, \\
\bar{\nu}_{\mathbb{l}} - \phi_{\mathbb{t}}\bar{\nu}_{\mathbb{t}} &= \bar{\kappa}\bar{\rho}_{O\mathbb{l}}, \qquad \bar{\nu}_{\mathbb{c}} - \phi_{\mathbb{v}}\bar{\nu}_{\mathbb{v}} = \bar{\kappa}\bar{\rho}_{O\mathbb{c}}, \qquad \bar{\nu}_{\mathbb{o}} - \phi_{\mathbb{u}}\bar{\nu}_{\mathbb{u}} = \bar{\kappa}\bar{\rho}_{O\mathbb{o}}, \\
\bar{\nu}_i &= \bar{\kappa}\bar{\rho}_{Ii} \quad \forall\; i \in \widehat{\mathbb{N}} \cup \{\mathbb{e}, \mathbb{l}, \mathbb{w}_1, \ldots, \mathbb{w}_w\}, \qquad \bar{\nu}_{\mathbb{t}} = \bar{\kappa}\bar{\rho}_{I\mathbb{ta}}, \\
\bar{\nu}_i + \alpha_i \bar{\nu}_{\mathbb{o}} + \beta_i \bar{\nu}_{\mathbb{R}} &= \bar{\kappa}\bar{\rho}_{Ii} - \bar{\gamma}_Y \psi_{\mathbb{o}} \alpha_i - \bar{\gamma}_C \psi_{\mathbb{R}} \beta_i \quad \forall\; i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \\
\bar{\nu}_{\mathbb{k}} + \alpha_{\mathbb{k}} \bar{\nu}_{\mathbb{o}} &= \bar{\kappa}\bar{\rho}_{I\mathbb{k}} - \bar{\gamma}_Y \psi_{\mathbb{o}} \alpha_{\mathbb{k}}.
\end{aligned}
\tag{23}
$$

Under the price normalisation rule that sets producer price of one output or input (say, price of labour, a good with index $\mathbb{L} \in \widehat{\mathbb{N}}$) equal to one, (23) implies that $\bar{\kappa} = \bar{\nu}_{\mathbb{L}}$; *i.e.*, if labour time is the numeraire commodity, then $\bar{\kappa}$ is its social shadow price. We will interpret the numeraire commodity as money.

The following theorem, whose proof can be found in the appendix, employs (23) to characterise the set of second-best optimal intermediate input taxes, which is denoted by Ω_η.

Theorem 6 *Suppose Assumptions 1 to 4 hold and we are at the second-best optimal configuration $\bar{\Upsilon}$. Define the following set of intermediate input tax vectors:*

$$
\Omega_\eta := \left\{ \eta \in \mathbf{R}^{\mathfrak{N}_O + 2} \;\middle|\; \eta_i = 0 \quad \forall\; i \in \widehat{\mathbb{N}} \cup \{\mathbb{ta}, \mathbb{r}\} \right.
$$
$$
\left. \wedge \quad \langle \eta_{\mathbb{n}}, \eta_{\mathbb{cn}}, \eta_{\mathbb{cw}}, \eta_{\mathbb{v}} \rangle \in \Lambda_{\mathbb{nwcv}} \quad \wedge \quad \langle \eta_{\mathbb{l}}, \eta_{\mathbb{tl}} \rangle \in \Lambda_{\mathbb{lt}} \quad \wedge \quad \langle \eta_{\mathbb{s}}, \eta_{\mathbb{b}}, \eta_{\mathbb{k}}, \eta_{\mathbb{o}}, \eta_{\mathbb{R}}, \eta_{\mathbb{u}} \rangle \in \Lambda_{\mathbb{sbkoRu}} \right\}
$$

where

- $\Lambda_{\mathbb{nwcv}}$ *is the set of vectors* $\langle \eta_{\mathbb{n}}, \eta_{\mathbb{w}}, \eta_{\mathbb{cn}}, \eta_{\mathbb{cw}}, \eta_{\mathbb{v}} \rangle \in \mathbf{R}^{2N+3}$ *that solve*

$$
\begin{aligned}
\eta_{\mathbb{n}} + \phi_{\mathbb{cn}} \eta_{\mathbb{cn}} + \phi_{\mathbb{v}} \phi_{\mathbb{cn}} \eta_{\mathbb{v}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cn}}, \\
\eta_{\mathbb{w}_i} + \phi_{\mathbb{cw}_i} \eta_{\mathbb{cw}_i} + \phi_{\mathbb{v}} \phi_{\mathbb{cw}_i} \eta_{\mathbb{v}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cw}_i} \quad \forall\; i = 1, \ldots, w.
\end{aligned}
\tag{24}
$$

- $\Lambda_{\mathbb{lt}}$ *is the set of vectors* $\langle \eta_{\mathbb{l}}, \eta_{\mathbb{tl}} \rangle \in \mathbf{R}^2$ *that solve*

$$
\eta_{\mathbb{l}} + \phi_{\mathbb{t}} \eta_{\mathbb{tl}} = 0. \tag{25}
$$

so that $\nabla_\rho^\top \mathfrak{y} \left[\sum_{i=\mathbb{s},\mathbb{b},\mathbb{k}} \bar{\sigma}^{Yi} + \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \bar{\sigma}^{Ci} + \bar{\sigma}^{\mathbb{co}} + \bar{\sigma}^{\mathbb{n}} \right] = \bar{\gamma}_C \nabla_\rho \mathfrak{z}_C + \bar{\gamma}_Y \nabla_\rho \mathfrak{z}_Y$.

- $\Lambda_{\mathbb{smbkoRu}}$ *is the set of vectors* $\langle \eta_{\mathbb{s}}, \eta_{\mathbb{m}}, \eta_{\mathbb{b}}, \eta_{\mathbb{k}}, \eta_{\mathbb{o}}, \eta_{\mathbb{R}}, \eta_{\mathbb{u}} \rangle \in \mathbf{R}^7$ *that solve*

$$\begin{aligned} \eta_i + \alpha_i \eta_{\mathbb{o}} + \beta_i \eta_{\mathbb{R}} + \alpha_i \phi_{\mathbb{u}} \eta_{\mathbb{u}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_i + \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} \beta_i \quad \forall\ i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \\ \eta_{\mathbb{k}} + \alpha_{\mathbb{k}} \eta_{\mathbb{o}} + \alpha_{\mathbb{k}} \phi_{\mathbb{u}} \eta_{\mathbb{u}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_{\mathbb{k}}. \end{aligned} \tag{26}$$

If $\bar{p} \in \mathcal{P}(\bar{\rho})$ *and* $\bar{\eta}$ *is the vector of intermediate input taxes associated with* $\bar{p}$ *then* $\bar{\eta} \in \Omega_\eta$*. Conversely, if* $\bar{\eta} \in \Omega_\eta$ *then there exists* $p \in \mathbf{R}^{\mathfrak{N}}$ *satisfying (5) such that* $\bar{\eta}$ *is the vector of intermediate input taxes associated with* p*. In particular, if* $p \in \mathbf{R}^{\mathfrak{N}}_+$*, then* $p \in \mathcal{P}(\bar{\rho})$*.*

3.2 A discussion of Theorem 6.

Two points regarding optimal intermediate input tax rates in the set Ω_η are worth noting:

Firstly, in contrast to DM, Theorem 6 shows that, in the presence of production externalities and non-substitutable inputs, optimal intermediate input taxes are not all zero. Rather, they play the role of Pigouvian taxes in controlling the generation of production externalities by

(i) directly taxing the sale of externality-causing goods[37] and/or

(ii) taxing the sale of non-substitutable inputs or the outputs produced by them that are located upstream or downstream relative to the location of the externality-causing good in the manufacturing chain that includes the externality-causing good [38] and/or

(ii) taxing the sale of inputs that are highly complementary to the use of the externality-causing goods as inputs in the production process.[39]

Secondly, the theorem shows that the optimal intermediate input tax vector is not unique, *i.e.,* the set of optimal intermediate input tax vectors Ω_η is not a singleton. This is because, the set of producer price vectors corresponding to the second-best optimal effective producer price vector $\bar{\rho}$ is $\mathcal{P}(\bar{\rho})$, which is not a singleton; and every producer price vector in $\mathcal{P}(\bar{\rho})$ is associated with a vector of intermediate input taxes that lies in Ω_η. Further, conditions (24) to (26) also imply that there are some restrictions and, at the same time, considerable degrees of freedom in choosing optimal intermediate input tax rates.

3.2.1 The manufacturing chain including coal.

To understand the above two aspects of Theorem 6, consider *e.g.,* the manufacturing chain that includes coal, which is an externality-causing good. As Figure 1 shows, the chain begins with the mining of coal from underneath the earth followed by its processing and sale to the thermal electricity generating sector and other sectors producing outputs such as cement, bricks, iron

[37] For example, by taxing coal, oil, or road service.

[38] To illustrate this, the manufacturing chain containing coal is discussed below.

[39] To illustrate this, the implication of road service and oil being complementary to the use vehicles by firms for motoring is discussed below.

and steel, paper and pulp *etc.*, denoted by $\mathbb{w}_1, \ldots, \mathbb{w}_w$. The figure shows various points at which the carbon externality can be taxed in the manufacturing chain that includes coal.

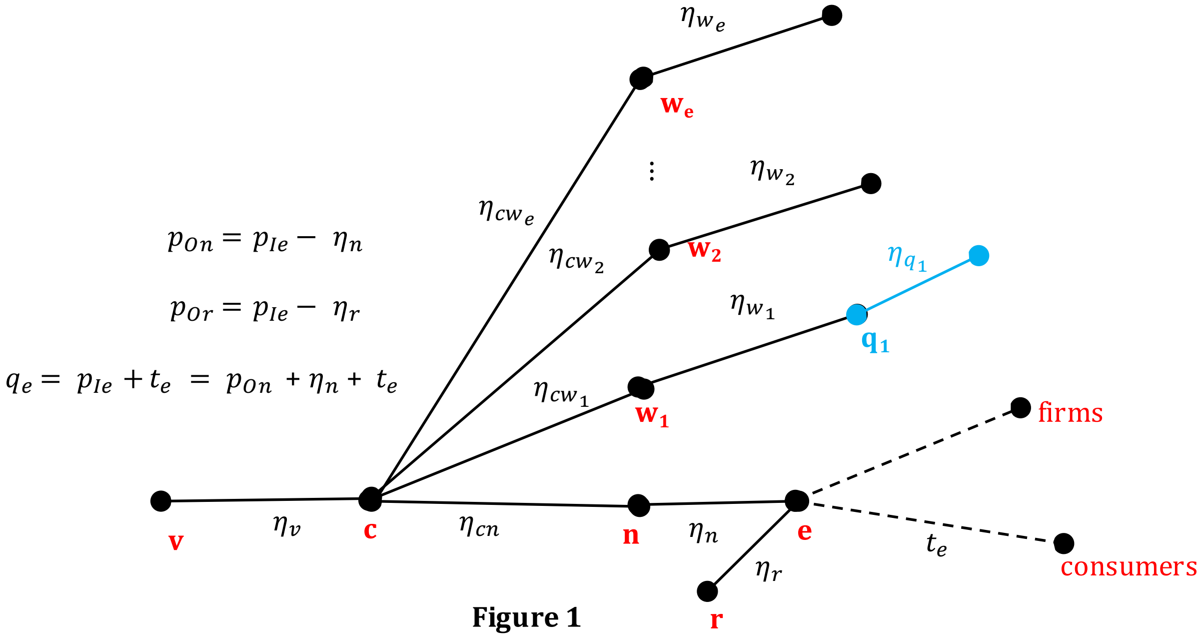

Figure 1

The link between various intermediate input tax rates in this manufacturing chain is described by (24). As will be discussed in Section 5, $\frac{\bar{\gamma}_Y}{\bar{\kappa}}$ is a money-metric measure of the social marginal cost (SMC) of the fossil-fuel (carbon) externality. Thus, the right-sides of the two conditions in (24) are, respectively, the money-metric measures of increases in the social costs due to the generation of an extra unit of thermal electricity and production of an extra unit of good $\mathbb{w}_i$ for $i = 1, \ldots, w$, all of which employ coal as an input.

For example, $\psi_{\mathbb{c}}\phi_{\mathbb{cn}}$ is the amount of CO_2 emission (*i.e.*, the carbon externality) generated by the amount of coal needed to produce one unit of thermal electricity (which is given by $\phi_{\mathbb{cn}}$). Hence, $\frac{\bar{\gamma}_j}{\bar{\kappa}}\psi_{\mathbb{c}}\phi_{\mathbb{cn}}$ is a money-metric measure of SMC of thermal electricity. Since coal is a non-substitutable input in the production of thermal electricity and $\phi_{\mathbb{cn}}$ amount of it is necessarily used to produce one unit of thermal electricity, $\frac{\bar{\gamma}_j}{\bar{\kappa}}\psi_{\mathbb{c}}\phi_{\mathbb{cn}}$ is also the money-metric measure of increase in social cost from using $\phi_{\mathbb{cn}}$ amount of coal for thermal electricity generation. Since $\phi_{\mathbb{v}}\phi_{\mathbb{cn}}$ amount of the output of the coal-mining sector is necessarily used to produce one unit of thermal electricity, $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{c}}\phi_{\mathbb{cn}}$ is also the money-metric measure of increase in social cost from using $\phi_{\mathbb{v}}\phi_{\mathbb{cn}}$ amount of the output of the coal-mining sector for thermal electricity generation. Externality correction requires ensuring that the revenue generated by Pigouvian taxation of a unit of thermal-electricity is equal to the monetised value of the SMC of thermal electricity. This revenue can be generated by (i) taxing sale of thermal electricity at the rate $\bar{\eta}_{\mathbb{n}}$ or (ii) taxing $\phi_{\mathbb{cn}}$ amount of processed coal sold to the the thermal electricity-generating sector at the rate $\bar{\eta}_{\mathbb{cn}}$ or (iii) taxing sale of $\phi_{\mathbb{v}}\phi_{\mathbb{cn}}$ amount of output of the coal-mining sector that finds its way, after processing, into the thermal electricity-generating sector at the rate $\bar{\eta}_{\mathbb{v}}$ or (iv) by distributing the Pigouvian tax burden, which is equal to the monetised value of the SMC of thermal electricity, between (i) to (iii) above. The first condition of (24) hence says that the tax rates $\bar{\eta}_{\mathbb{n}}$, $\bar{\eta}_{\mathbb{cn}}$, and $\bar{\eta}_{\mathbb{v}}$ should be chosen such that the sum of intermediate input tax revenues collected from the thermal electricity-producing sector, the coal-processing sector, and the coal-mining sector when one unit of thermal electricity is produced is equal to the monetised value of the SMC of thermal electricity.

Similarly, production of good $\mathbb{w}_i$ requires $\phi_{\mathbb{cw}_i}$ amount of processed coal, which in turn requires $\phi_{\mathbb{v}}\phi_{\mathbb{cw}_i}$ amount of mined coal. Hence, the second condition of (24) says that the tax rates $\bar{\eta}_{\mathbb{w}_i}$, $\bar{\eta}_{\mathbb{cw}_i}$, and $\bar{\eta}_{\mathbb{v}}$ should be chosen such that the sum of intermediate input tax revenues collected from the sector that produces good $\mathbb{w}_i$, the coal-processing sector, and the coal-mining sector when one unit of good $\mathbb{w}_i$ is produced is equal to the monetised value of the SMC of good $\mathbb{w}_i$. This is true for all $i = 1, \ldots, w$.

In fact, if the manufacturing chain including coal is longer, for example, if $\mathbb{w}_1$ is a non-substitutable input in the production of a new good $\mathbb{q}$ with $\phi_{\mathbb{w}_1\mathbb{q}}$ being the amount of good $\mathbb{w}_1$ required to produce one unit of good $\mathbb{q}$,[40] then a part of the burden of the Pigouvian tax can also be passed on to sector $\mathbb{q}$, *i.e.,* the tax rates $\bar{\eta}_{\mathbb{q}}$, $\bar{\eta}_{\mathbb{w}_1}$, $\bar{\eta}_{\mathbb{cw}_1}$, and $\bar{\eta}_{\mathbb{v}}$ should be chosen to satisfy

$$\bar{\eta}_{\mathbb{q}} + \phi_{\mathbb{w}_1\mathbb{q}}\bar{\eta}_{\mathbb{w}_1} + \phi_{\mathbb{w}_1\mathbb{q}}\phi_{\mathbb{cw}_1}\bar{\eta}_{\mathbb{cw}_1} + \phi_{\mathbb{w}_1\mathbb{q}}\phi_{\mathbb{cw}_1}\phi_{\mathbb{v}}\bar{\eta}_{\mathbb{v}} = \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{c}}\phi_{\mathbb{cw}_1}\phi_{\mathbb{w}_1\mathbb{q}}.$$

Policy-oriented literature often discusses the welfare-equivalence of upstream or downstream taxation of externality along the manufacturing chain, while pointing out the administrative efficacy of upstream externality taxation. For example, Metcalf (2009) argues that, for administrative ease, a carbon tax should be levied upstream on fuel producers rather than downstream on fuel users. This is because one expects the former taxable units to be fewer in number than the latter. Thus, tax should be applied at the mine mouths for domestic coal and either on the crude oil as it enters the refineries or on sale of various products produced by the refineries. Similar arguments in favour of upstream externality taxation are also made by Fullerton, Leicester, and Smith (2010), Mansur (2012), and Barthold (1994). This work hence provides a theoretical foundation for such arguments.

Theorem 6 also demonstrates that the welfare equivalence of upstream and downstream externality taxation needs to be qualified. This is because (24) can be viewed as a system of $w + 1$ equations in $2w + 3$ unknowns, namely, $\bar{\eta}_{\mathbb{w}_i}$ and $\bar{\eta}_{\mathbb{cw}_i}$ for all $i = 1, \ldots, w$; $\bar{\eta}_{\mathbb{n}}$; $\bar{\eta}_{\mathbb{cn}}$; and $\bar{\eta}_{\mathbb{v}}$. Hence, there are effectively $w + 2$ degrees of freedom in choosing these tax rates at the social optimum $\bar{\Upsilon}$. For example, (24) shows that $w + 1$ tax rates $\bar{\eta}_{\mathbb{n}}$, $\bar{\eta}_{\mathbb{w}_1}, \ldots, \bar{\eta}_{\mathbb{w}_w}$ can be expressed as functions of the remaining $w + 2$ tax rates $\bar{\eta}_{\mathbb{cn}}$, $\bar{\eta}_{\mathbb{v}}$, $\bar{\eta}_{\mathbb{cw}_1}, \ldots, \bar{\eta}_{\mathbb{cw}_w}$.

Fullerton, Leicester, and Smith (2010) discuss the climate change levy (CCL) in U.K., which is a single stage excise tax levied on industrial and commercial energy use of gas, electricity,

[40] For example, cement can be considered a non-substitutable input in the construction industry or steel is a non-substitutable input in the production of steel utensils.

and coal. A single tax rate on the sale of coal to different uses, such as to the thermal energy producing sector or cement and steel industries, is employed. Hence, this can be interpreted as a case where $\bar{\eta}_{\mathbb{cn}} = \bar{\eta}_{\mathbb{cw}_i} \equiv \bar{\eta}_{\mathbb{c}}$ for $i = 1, \ldots, w$. Suppose no tax is imposed on coal-mining, *i.e.*, $\bar{\eta}_{\mathbb{v}} = 0$. In this case, (24) implies

$$\begin{aligned}
\bar{\eta}_{\mathbb{n}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cn}} - \phi_{\mathbb{cn}} \bar{\eta}_{\mathbb{c}} \\
\bar{\eta}_{\mathbb{w}_i} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cw}_i} - \phi_{\mathbb{cw}_i} \bar{\eta}_{\mathbb{c}} \quad \forall\, i = 1, \ldots, w.
\end{aligned}$$

Hence, the above shows that, if $\bar{\eta}_{\mathbb{c}}$ is chosen so that there is a positive tax on thermal electricity (as is done under CCL), then, in general, the second-best tax structure will also require non-zero intermediate input taxation/subsidization of other sectors such as steel and cement, which use coal as a non-substitutable input.

Alternatively, to correct for the carbon externality from coal, it is enough to tax just coal either at the processing stage or at the mining-stage, *i.e.*, set $\bar{\eta}_{\mathbb{n}} = \bar{\eta}_{\mathbb{w}_i} = 0$ for all $i = 1, \ldots, w$ combined with

- *either* $\bar{\eta}_{\mathbb{cn}} = \bar{\eta}_{\mathbb{cw}_i} \equiv \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}}$ for $i = 1, \ldots, w$
- *or* $\bar{\eta}_{\mathbb{cn}} = \bar{\eta}_{\mathbb{cw}_i} = 0$ for $i = 1, \ldots, w$ and $\bar{\eta}_{\mathbb{v}} = \frac{\bar{\gamma}_Y}{\bar{\kappa}} \frac{\psi_{\mathbb{c}}}{\phi_{\mathbb{v}}}$

As opposed to the CCL in U.K., a coal excise is levied on mined coal in the U.S. and India. The coal excise rates in the U.S. depend on the level at which the mining takes place – surface or underground (see Metcalf (2009)).

3.2.2 Oil, road service, and vehicle use as complementary inputs required for motoring.

When firms use the motoring services of vehicles as inputs for transportation, then they also need the input services of roads and a motoring fuel such as oil. But the use of these services leads to the generation of both carbon and congestion externalities. Figure 2 below shows various points of taxation of carbon and congestion externality generated when motoring service from vehicle of size i is employed.

Thus, (26) shows that congestion and carbon externalities from motoring that are generated by firms can be jointly controlled by intermediate input taxation of oil ($\bar{\eta}_{\mathbb{o}}$) or motoring services from vehicles and road ($\bar{\eta}_i$ for $i = \mathbb{s}, \mathbb{b}, \mathbb{m}$ and $\bar{\eta}_{\mathbb{R}}$) or other non-motoring services that require use of oil ($\bar{\eta}_{\mathbb{k}}$). The right-sides of (26) are the monetised values of the SMCs of services from big, medium, and small-sized vehicles and the monetised value of SMC of non-motoring services that also require oil. In particular, for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$, the former is the sum of monetised value of social costs of carbon and congestion externalities generated by a unit of service from the i^{th} vehicle, which requires α_i amount of oil and β_i amount of road services. Hence, the complementarities between (or the need to simultaneously use) services from vehicle, roads, and oil as well as the non-substitutability of crude of oil in producing refined oil imply that $\frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_i + \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} \beta_i$ is also equal to (i) the monetised value of social cost of using α_i amount of oil to drive the i^{th} vehicle-type, (ii) the monetised value of social cost of using β_i amount of road service to drive the i^{th} vehicle-type, and (iii) the monetised value of social cost from using $\alpha_i \phi_{\mathbb{u}}$ amount of crude oil that is required to produce α_i amount of (refined) oil to drive the i^{th} vehicle-type.

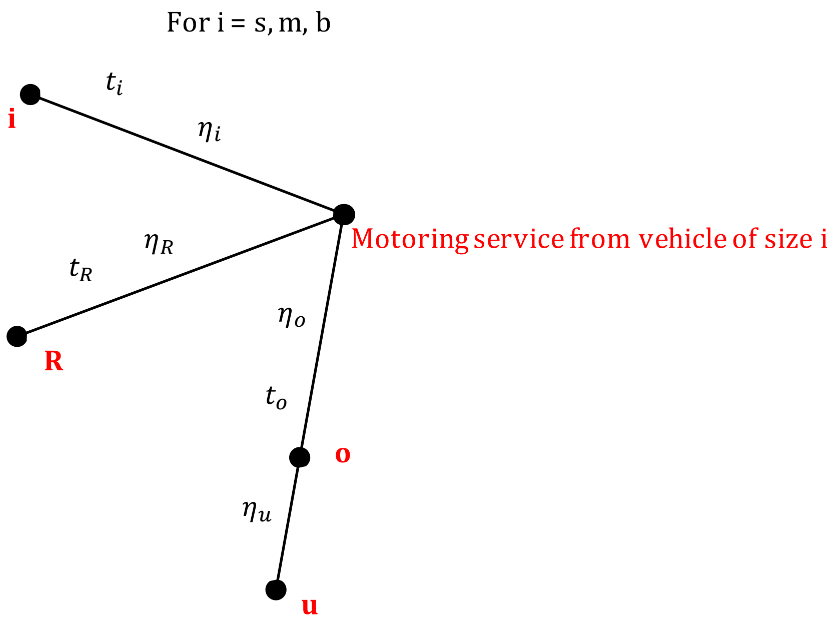

Figure 2

Thus, the conditions in (26) show that the correction of externalities of carbon and congestion from use of oil and road services requires choosing intermediate input tax rates $\bar{\eta}_i$ for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$; $\bar{\eta}_{\mathbb{o}}$; $\bar{\eta}_{\mathbb{R}}$; and $\bar{\eta}_{\mathbb{u}}$ such that the sum of the Pigouvian tax revenues from taxing good i, oil, road, and unrefined oil when one unit of the service from good i is used is equal to the monetised value of its SMC.

Further (26) consists of four equations in seven unknowns, namely, $\bar{\eta}_i$ for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$; $\bar{\eta}_{\mathbb{o}}$; $\bar{\eta}_{\mathbb{R}}$; and $\bar{\eta}_{\mathbb{u}}$. Hence, there are three degrees of freedom in choosing these tax rates at the second-best optimum $\bar{\Upsilon}$. For example, the optimal tax rates for vehicle use ($\mathbb{s}, \mathbb{m}, \mathbb{b}$) and non-motoring service from oil ($\mathbb{k}$) are determined once tax rates $\bar{\eta}_{\mathbb{o}}$, $\bar{\eta}_{\mathbb{R}}$, and $\bar{\eta}_{\mathbb{u}}$ are pegged. Moreover, (26) shows that motoring externalities cannot be corrected by only oil and road service taxation when there are vehicular differences in the extents to which carbon and congestion externalities are generated. This is because, when $\bar{\eta}_i = 0$ for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$, then the system (26) is over-determined (has four equations in three unknowns, $\bar{\eta}_{\mathbb{o}}$, $\bar{\eta}_{\mathbb{R}}$, and $\bar{\eta}_{\mathbb{u}}$) and, in general, will have no solutions. The system (26) also shows that, in principle, differential taxation of $\mathbb{s}$, $\mathbb{m}$, $\mathbb{b}$, and $\mathbb{k}$ alone (with no taxation of oil or road service) is enough to completely correct the externalities from congestion and carbon. This is true when we choose these tax rates as

$$\bar{\eta}_i \;=\; \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{o}}\alpha_i + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_i \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b}; \qquad \eta_{\mathbb{k}} \;=\; \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{o}}\alpha_{\mathbb{k}}.$$

Excise taxation of fuel for motoring is widely employed, *e.g.,* motor fuel excise tax on gasoline in the U.S., fuel duties on unleaded petrol and diesel in the U.K., and central excise on petrol and diesel in India.[41] In addition excise taxes and duties that vary with vehicle types are also in place, *e.g.,* the gas guzzler tax in U.S., which differentially taxes vehicles on the basis of their fuel mileage (see, Metcalf (2008)); the vehicle excise duty (VED) in U.K., which is an annual charge levied on the owners of vehicles based on CO_2 emissions per kilometres; and the differential rates of compensatory-cesses (in addition to GST) levied in India on cars based on their size (length, fuel-type used, and carrying capacity). Levell, O'Connell, and Smith (2016) argue that the effectiveness of motoring excises depends on the extent to which they allow tax payments to increase with increase in the vehicle use. In this regards they argue that the VED is inferior to the fuel excise, as the former is collected as a lump-sum amount every year, while in the latter case, tax payments vary with fuel usage, which will in turn can be expected to vary with vehicle use. However, while fuel excise is effective in penalising vehicle usage for CO_2 emissions, it is not so effective for controlling congestion externalities as it does not distinguish between vehicular differences in road service (*e.g.,* road-space) requirement while motoring. They argue that a road-tax is an effective means of controlling congestion externalities.

In our stylised model, tax payments vary depending on the extent of usage of vehicular services, measured *e.g.,* in terms of kilometres motored and the tax rates vary across vehicles on the basis of their size, mileage (fuel per unit of service), and road services requirements per unit of vehicular service.

3.2.3 On intermediate input taxation of consumption-externality causing goods.

In our model the use or production of consumption-externality causing goods $\mathbb{t}$ and $\mathbb{l}$ by firms generate no externalitis. Hence, the right-side of (25), which is the SMC of producing good $\mathbb{l}$ is zero. Since $\phi_{\mathbb{t}}$ amount of good $\mathbb{t}$ is required to produce one unit of good $\mathbb{l}$, the above also implies that the social cost of using $\phi_{\mathbb{t}}$ amount of good $\mathbb{t}$ while producing good $\mathbb{l}$ is also zero. Hence, condition (25) says that intermediate input tax rates $\bar{\eta}_{\mathbb{l}}$ and $\bar{\eta}_{\mathbb{tl}}$ should be chosen so that the sum of intermediate-input tax revenues collected from sectors $\mathbb{t}$ and $\mathbb{l}$ when one unit of good $\mathbb{l}$ is produced should be zero. Note that condition (25) is consistent with no intermediate input taxation of both goods, *i.e.,* with $\bar{\eta}_{\mathbb{l}} = \bar{\eta}_{\mathbb{tl}} = 0$. But it is also consistent with taxation of one of the two goods and the subsidisation of the other.

3.2.4 No intermediate input taxation of remaining goods and an implicit subsidy on renewables.

Since goods in $\widehat{\mathbb{N}} \cup \{\mathbb{to}, \mathbb{r}\}$ do not lie in the manufacturing chain of goods that include the externality-causing goods nor are they complementary to externality generating goods in their role as inputs, Theorem 6 shows that they are exempted from intermediate input taxation under

[41] The central excise on petrol and diesel seems not to be an environmental levy in India. A VAT at the state level is also levied on the price inclusive of the central excise. In addition, a small pollution cess with surcharge is also imposed.

the optimal commodity tax structure. Thus, grapes purchased for producing non-alcoholic foods should not be taxed, while it can be taxed if it is purchased for producing wine.

It is to be noticed that there is an implicit subsidy on production of electricity using renewables at the second-best optimum. This is because, Theorem 6 shows that the intermediate input tax on renewable energy $\bar{\eta}_{\mathbb{r}}$ is zero (as renewable energy production generates no externalities), while it can be chosen to be positive on non-renewable energy, *i.e.,* (24) shows that $\bar{\eta}_{\mathbb{n}}$ can be chosen to be positive.[42] Hence, given that all firms face a common input price of electricity $p_{I\mathbb{e}}$, the price received by renewable electricity sellers $p_{O\mathbb{r}}$ (which is equal to $p_{I\mathbb{e}} - \bar{\eta}_{\mathbb{r}} = p_{I\mathbb{e}}$) is higher than the price $p_{O\mathbb{n}} = p_{I\mathbb{e}} - \bar{\eta}_{\mathbb{n}}$ received by sellers of thermal electricity.

4 Optimal consumption tax rates.

In this section, we will characterise and discuss the second-best optimal consumption tax rates and show that these can be decomposed into MPRR and externality-correction components. The MPRR component is redundant for some commodities such as oil, road service, and electricity; while the externality-correction components are absent for commodities that are not associated with externality generation.

4.1 Characterisation of second-best optimal consumption tax rates.

The FOCs with respect to effective consumer prices can be expressed in matrix format as:[43]

$$\begin{aligned} -\sum_h \mu^h \lambda^h \bar{\chi}^{h\top} - \sum_{i \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}, \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}\}} \bar{\nu}_i \sum_h \nabla_\delta x_i^h - \sum_{i = \mathbb{s}, \mathbb{m}, \mathbb{b}} [\bar{\nu}_i + \alpha_i \bar{\nu}_{\mathbb{o}} + \beta_i \bar{\nu}_{\mathbb{R}}] \sum_h \nabla_\delta x_i^h \\ - [\bar{\nu}_{\mathbb{k}} + \alpha_{\mathbb{k}} \bar{\nu}_{\mathbb{o}}] \sum_h \nabla_\delta x_{\mathbb{k}}^h - \bar{\gamma}_C \nabla_\delta^\top \mathcal{Z}_C - \bar{\gamma}_X \nabla_\delta^\top \mathcal{Z}_X - \bar{\gamma}_Y \nabla_\delta^\top \mathcal{Z}_Y = 0, \end{aligned}$$

Recalling the definition of χ^h in Section 2.3 and employing (23), (5), and (11) the above can be re-written as[44]

$$\sum_h \nabla_\delta^\top \chi^h \bar{\varrho} \;=\; \bar{b}, \tag{27}$$

[42]This will depend on how other intermediate input tax rates along the manufacturing chain that includes coal are chosen.

[43]Note that in writing the FOCs, we have employed (9). The Jacobian of consumer demand $\sum_h \nabla_\delta x^h$ is evaluated at the second-best optimal configuration $\bar{\Upsilon}$ and $\bar{x}^h$ is the consumption bundle of consumer $h = 1, \ldots, H$ at the second-best optimum corresponding to $\bar{\Upsilon}$.

[44]This follows because (23) and (5) imply that $\bar{\nu}_i = \bar{\kappa}\bar{p}_{Oi}$ for all $i \in \widetilde{\mathcal{N}} \cup \{\mathbb{t}\}$; $\bar{\nu}_{\mathbb{e}} = \bar{\kappa}\bar{p}_{O\mathbb{r}} = \bar{\kappa}\bar{p}_{I\mathbb{e}}$; $\bar{\nu}_{\mathbb{w}_i} = \bar{\kappa} p_{O\mathbb{w}_i} - \bar{\kappa}\phi_{\mathbb{c}\mathbb{w}_i}\bar{\eta}_{\mathbb{c}\mathbb{w}_i} - \bar{\kappa}\phi_{\mathbb{c}\mathbb{w}_i}\phi_{\mathbb{v}}\bar{\eta}_{\mathbb{v}} + \bar{\gamma}\psi_{\mathbb{c}}\phi_{\mathbb{c}\mathbb{w}_i}$ for all $i = 1, \ldots, w$; $\bar{\nu}_{\mathbb{l}} = \bar{\kappa}\bar{p}_{O\mathbb{l}} - \bar{\kappa}\phi_{\mathbb{t}}\bar{\eta}_{\mathbb{t}\mathbb{l}}$; $\bar{\nu}_{\mathbb{c}} = \bar{\kappa}\bar{p}_{O\mathbb{c}} - \bar{\kappa}\phi_{\mathbb{v}}\bar{\eta}_{\mathbb{v}}$; and $\bar{\nu}_{\mathbb{o}} = \bar{\kappa}\bar{p}_{O\mathbb{o}} - \bar{\kappa}\phi_{\mathbb{u}}\bar{\eta}_{\mathbb{u}}$. We can then employ (11) to express producer prices in terms of consumer prices q and consumption tax rates t and employ (10) to express q in terms of δ. Finally, we note that the differentiation of the effective budget constraint of any consumer h with respect to effective consumer prices yields $\sum_{i \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\}} \delta_i \frac{\partial x_i^h}{\partial \delta_j} = -x_j^h$ for all $j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\}$.

where

$$
\begin{aligned}
\bar{b} &:= \frac{\sum_h \mu^h \lambda^h \bar{\chi}}{\bar{\kappa}} - \bar{\chi}^h + \frac{\bar{\gamma}_C}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_C + \frac{\bar{\gamma}_X}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_X + \frac{\bar{\gamma}_Y}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_Y \\
\bar{\varrho} &:= \Big\langle \bar{t}_{\widehat{\mathbb{N}}},\ \bar{t}_{\mathbb{s}} + \alpha_{\mathbb{s}} [\bar{t}_{\mathbb{o}} + \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}}] + \beta_{\mathbb{s}} \bar{t}_{\mathbb{R}},\ \bar{t}_{\mathbb{m}} + \alpha_{\mathbb{m}} [\bar{t}_{\mathbb{o}} + \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}}] + \beta_{\mathbb{m}} \bar{t}_{\mathbb{R}},\ \bar{t}_{\mathbb{b}} + \alpha_{\mathbb{b}} [\bar{t}_{\mathbb{o}} + \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}}] + \beta_{\mathbb{b}} \bar{t}_{\mathbb{R}}, \\
&\quad \bar{t}_{\mathbb{k}} + \alpha_{\mathbb{k}} [\bar{t}_{\mathbb{o}} + \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}}],\ \bar{t}_{\mathbb{w}_1} + \phi_{\mathbb{cw}_1} \bar{\eta}_{\mathbb{cw}_1} + \phi_{\mathbb{cw}_1} \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} - \frac{\bar{\gamma}}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cw}_1},\ \ldots, \\
&\quad \bar{t}_{\mathbb{w}_w} + \phi_{\mathbb{cw}_w} \bar{\eta}_{\mathbb{cw}_w} + \phi_{\mathbb{cw}_w} \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} - \frac{\bar{\gamma}}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cw}_w},\ \bar{t}_{\mathbb{t}},\ \bar{t}_{\mathbb{e}},\ \bar{t}_{\mathbb{l}} + \phi_{\mathbb{t}} \bar{\eta}_{\mathbb{tl}},\ \bar{t}_{\mathbb{c}} + \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} \Big\rangle \\
&\equiv \Big\langle \bar{\varrho}_{\widehat{\mathbb{N}}},\ \bar{\varrho}_{\mathbb{s}},\ \bar{\varrho}_{\mathbb{m}},\ \bar{\varrho}_{\mathbb{b}},\ \bar{\varrho}_{\mathbb{k}},\ \bar{\varrho}_{\mathbb{w}_1},\ \ldots,\ \bar{\varrho}_{\mathbb{w}_w},\ \bar{\varrho}_{\mathbb{t}},\ \bar{\varrho}_{\mathbb{e}},\ \bar{\varrho}_{\mathbb{l}},\ \bar{\varrho}_{\mathbb{c}} \Big\rangle \in \mathbf{R}^{N-2} \qquad (28)
\end{aligned}
$$

For every $h = 1, \ldots, H$, $\nabla_\delta^\top \chi^h$ is the following $(N-2) \times (N-2)$-dimensional matrix that is evaluated at the second-best optimal configuration $\bar{\Upsilon}$.

$$
\nabla_\delta^\top \chi^h = \Big[\nabla_\delta x^h_{\widehat{\mathbb{N}}} \quad \nabla_\delta x^h_{\mathbb{s}} \quad \nabla_\delta x^h_{\mathbb{m}} \quad \nabla_\delta x^h_{\mathbb{b}} \quad \nabla_\delta x^h_{\mathbb{k}} \quad \nabla_\delta x^h_{\mathbb{w}} \quad \nabla_\delta x^h_{\mathbb{t}} \quad \nabla_\delta x^h_{\mathbb{e}} \quad \nabla_\delta x^h_{\mathbb{l}} \quad \nabla_\delta x^h_{\mathbb{c}} \Big]. \qquad (29)
$$

Assumption 7 *Rank of the* $(N-2) \times (N-2)$*-dimensional matrix* $\sum_h \nabla_\delta^\top \chi^h$ *is* $N-2$*, i.e., the matrix* $\sum_h \nabla_\delta^\top \chi^h$ *is full-ranked.*[45]

Noting that

$$
\begin{aligned}
\nabla_\delta \mathcal{Z}_Y &= \psi_{\mathbb{o}} \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i \sum_h \nabla_\delta x_i^h + \psi_{\mathbb{c}} \sum_h \nabla_\delta x_{\mathbb{c}}^h \\
\nabla_\delta \mathcal{Z}_X &= \psi_{\mathbb{l}} \sum_h \nabla_\delta x_{\mathbb{l}}^h + \psi_{\mathbb{t}} \sum_h \nabla_\delta x_{\mathbb{t}}^h \\
\nabla_\delta \mathcal{Z}_C &= \psi_{\mathbb{R}} \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i \sum_h \nabla_\delta x_i^h, \qquad (30)
\end{aligned}
$$

we find that there exist vectors ς^{Ci} for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$; ς^{Yi} for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$; $\varsigma^{Y\mathbb{c}}$; and ς^{Xi} for $i = \mathbb{l}, \mathbb{t}$ all in $\mathbf{R}^{N-2}$ such that the following hold

$$
\begin{aligned}
\sum_h \nabla_\delta^\top \chi^h \varsigma^{Ci} &= \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} \beta_i \sum_h \nabla_\delta x_i^h \ \forall\ i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \\
\sum_h \nabla_\delta^\top \chi^h \varsigma^{Yi} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_i \sum_h \nabla_\delta x_i^h \ \forall\ i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}, \\
\sum_h \nabla_\delta^\top \chi^h \varsigma^{Y\mathbb{c}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \sum_h \nabla_\delta^\top x_{\mathbb{c}}^h, \\
\sum_h \nabla_\delta^\top \chi^h \varsigma^{Xi} &= \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_i \sum_h \nabla_\delta^\top x_i^h \ \forall\ i = \mathbb{l}, \mathbb{t}.
\end{aligned}
$$

This is because the structures of matrix $\sum_h \nabla_\delta^\top \chi^h$ and the gradients $\nabla_\delta \mathcal{Z}_Y$, $\nabla_\delta \mathcal{Z}_X$, and $\nabla_\delta \mathcal{Z}_X$

[45]The assumption that the Jacobian of consumer demands with respect to prices is full ranked is often made in optimal tax literature. See for example Sandmo (1975).

as seen in (29) and (30) imply that we can choose

$$\begin{array}{rclccl}
\varsigma_j^{Ci} & = & 0 \quad \forall\, j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}, i\} & \text{and} & \varsigma_i^{Ci} = \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} \beta_i & \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b} \\
\varsigma_j^{Yi} & = & 0 \quad \forall\, j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}, i\} & \text{and} & \varsigma_i^{Yi} = \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_i & \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k} \\
\varsigma_j^{Y\mathbb{c}} & = & 0 \quad \forall\, j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}, \mathbb{c}\} & \text{and} & \varsigma_{\mathbb{c}}^{Y\mathbb{c}} = \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} & \\
\varsigma_j^{Xi} & = & 0 \quad \forall\, j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}, i\} & \text{and} & \varsigma_i^{Xi} = \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_i & \forall\, i = \mathbb{l}, \mathbb{t}.
\end{array} \tag{31}$$

Hence, we have

$$\sum_h \nabla_\delta^\top \chi^h \left[\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \varsigma^{Ci} + \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \varsigma^{Yi} + \varsigma^{Y\mathbb{c}} + \sum_{i=\mathbb{l},\mathbb{t}} \varsigma^{Xi} \right] = \frac{\bar{\gamma}_C}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_C + \frac{\bar{\gamma}_X}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_X + \frac{\bar{\gamma}_Y}{\bar{\kappa}} \nabla_\delta \mathcal{Z}_Y. \tag{32}$$

Assumption 7 and (27) imply that $\bar{\varrho}$ is a *unique* solution to the linear system of equations $\sum_h \nabla_\delta^\top \chi^h \varrho = \bar{b}$. Given the structure of $\bar{b}$ (see (28)), (32) implies that there hence exists a unique vector $\bar{\mathbf{t}}^{MPRR}$ that solves

$$\bar{\varrho} = \bar{\mathbf{t}}^{MPRR} + \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \varsigma^{Ci} + \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \varsigma^{Yi} + \varsigma^{Y\mathbb{c}} + \xi^{X\mathbb{l}} \tag{33}$$

such that

$$\sum_h \nabla_\delta^\top \chi^h \bar{\mathbf{t}}^{MPRR} = \frac{\sum_h \mu^h \lambda^h \bar{\chi}^h}{\bar{\kappa}} - \sum_h \bar{\chi}^h. \tag{34}$$

Hence, (33) and (31) imply that

$$\begin{array}{rcl}
\bar{\varrho}_i & = & \bar{\mathbf{t}}_i^{MPRR} \quad \forall\, i \in \widehat{\mathcal{N}} \cup \{\mathbb{t}, \mathbb{e}, \mathbb{w}_1, \ldots, \mathbb{w}_w\} \\
\bar{\varrho}_i & = & \bar{\mathbf{t}}_i^{MPRR} + \frac{\bar{\gamma}_C}{\bar{\kappa}} \beta_i + \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_i \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b} \\
\bar{\varrho}_{\mathbb{k}} & = & \bar{\mathbf{t}}_{\mathbb{k}}^{MPRR} + \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \alpha_{\mathbb{k}} \\
\bar{\varrho}_{\mathbb{c}} & = & \bar{\mathbf{t}}_{\mathbb{c}}^{MPRR} + \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \\
\bar{\varrho}_i & = & \bar{\mathbf{t}}_i^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_i \quad \forall\, i = \mathbb{l}, \mathbb{t}
\end{array} \tag{35}$$

Given $\bar{\eta} \in \Omega_\eta$, the following theorem identifies the set of optimal consumption tax vectors at the second-best optimum $\bar{\Upsilon}$. The proof follows from employing (28) to substitute for $\bar{\varrho}$ in (35).

Theorem 8 *Suppose Assumptions 1 to 4 and 7 hold and we are at the second-best optimum $\bar{\Upsilon}$. Given $\bar{\eta} \in \Omega_\eta$, define the following set of consumption tax vectors:*

$$\Omega_t(\bar{\eta}) = \left\{ \bar{t} \in \mathbf{R}^N \;\middle|\; \bar{t}_i = \bar{\mathbf{t}}_i^{MPRR} \quad \forall\, i \in \widehat{\mathcal{N}} \cup \{\mathbb{e}\} \;\wedge\; \langle \bar{t}_{\mathbb{s}}, \bar{t}_{\mathbb{m}}, \bar{t}_{\mathbb{b}}, \bar{t}_k, \bar{t}_{\mathbb{R}}, \bar{t}_{\mathbb{o}} \rangle \in \Xi_{\mathbb{sm},\mathbb{bk}}\left(\bar{\eta}_{\mathbb{u}}\right) \wedge \right.$$
$$\left. \bar{t}_{\mathbb{w}} \in \Xi_{\mathbb{w}}\left(\bar{\eta}_{\mathbb{w}}, \bar{\eta}_{\mathbb{cw}} \bar{\eta}_{\mathbb{v}}\right) \;\wedge\; \bar{t}_{\mathbb{l}} \in \Xi_{\mathbb{l}}\left(\bar{\eta}_{\mathbb{l}}\right) \;\wedge\; \bar{t}_{\mathbb{t}} \in \Xi_{\mathbb{t}} \;\wedge\; \bar{t}_{\mathbb{c}} \in \Xi_{\mathbb{c}}\left(\bar{\eta}_{\mathbb{v}}\right) \right\}$$

where

- $\Xi_{\mathbb{smbkRo}}\left(\bar{\eta}_{\mathbb{u}}\right)$ *is the set of vectors* $\langle t_{\mathbb{s}}, t_{\mathbb{m}}, t_{\mathbb{b}}, t_k, t_{\mathbb{R}}, t_{\mathbb{o}} \rangle \in \mathbf{R}^6$ *that solve*

$$\begin{aligned} t_i &= \bar{\mathbf{t}}_i^{MPRR} + \beta_i \left[\frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} - t_{\mathbb{R}} \right] + \alpha_i \left[\frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} - t_{\mathbb{o}} - \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}} \right], \;\forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b} \\ t_{\mathbb{k}} &= \bar{\mathbf{t}}_{\mathbb{k}}^{MPRR} + \alpha_i \left[\frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} - t_{\mathbb{o}} - \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}} \right] \end{aligned} \tag{36}$$

- $\Xi_{\mathbb{w}}\left(\bar{\eta}_{\mathbb{w}}, \bar{\eta}_{\mathbb{cw}} \bar{\eta}_{\mathbb{v}}\right)$ *is the set of vectors* $t_{\mathbb{w}} \in \mathbf{R}^w$ *that solve*

$$t_{\mathbb{w}_i} = \bar{\mathbf{t}}_{\mathbb{w}_i}^{MPRR} + \frac{\bar{\gamma}}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cw}_i} - \phi_{\mathbb{cw}_i} \bar{\eta}_{\mathbb{cw}_i} - \phi_{\mathbb{cw}_i} \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} \quad \forall\, i = 1, \ldots, w \tag{37}$$

- $\Xi_{\mathbb{l}}\left(\bar{\eta}_{\mathbb{l}}\right)$ *is the set of scalars* $t_{\mathbb{l}} \in \mathbf{R}$ *that solve*

$$t_{\mathbb{l}} = \bar{\mathbf{t}}_{\mathbb{l}}^{MPRR} + \left[\frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{l}} - \phi_{\mathbb{t}} \bar{\eta}_{\mathbb{tl}} \right] \tag{38}$$

- $\Xi_{\mathbb{t}}$ *is the set of scalars* $t_{\mathbb{t}} \in \mathbf{R}$ *that solve*

$$t_{\mathbb{t}} = \bar{\mathbf{t}}_{\mathbb{l}}^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{t}} \tag{39}$$

- $\Xi_{\mathbb{c}}\left(\bar{\eta}_{\mathbb{v}}\right)$ *is the set of scalars* $t_{\mathbb{c}} \in \mathbf{R}$ *that solve*

$$t_{\mathbb{c}} = \bar{\mathbf{t}}_{\mathbb{c}}^{MPRR} + \left[\frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} - \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} \right] \tag{40}$$

Suppose $\bar{p} \in \mathcal{P}\left(\bar{\rho}\right)$, $\bar{q} \in \mathcal{Q}\left(\bar{\delta}\right)$, $\bar{\eta}$ *is the vector of intermediate input taxes associated with* $\bar{p}$, *and* $\bar{t}$ *is the vector of consumption taxes associated with* $\bar{q}$ *and* $\bar{p}$. *Then* $\bar{t} \in \Omega_t\left(\bar{\eta}\right)$.
Conversely, suppose $t \in \Omega_t\left(\bar{\eta}\right)$ *and there exists* $\bar{p} \in \mathcal{P}\left(\bar{\rho}\right)$ *such that* $\bar{\eta}$ *is the vector of intermediate input taxes associated with* $\bar{p}$. *If* q *is derived from* t *and* $\bar{p}$ *using (11) and* $q \in \mathbf{R}_+^N$, *then* $q \in \mathcal{Q}\left(\bar{\delta}\right)$.

4.2 The MPRR and externality excise components of consumption taxes.

The vector of wedges t between the consumer prices q and producer output prices p_O can be decomposed into a non-externality and an externality component:

$$\begin{aligned} t_i &= \mathbf{t}_i^{MPRR} + \left(t_i - \mathbf{t}_i^{MPRR}\right) \quad \forall\, i \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\} \\ &=: \mathbf{t}_i^{MPRR} + t_i^{\mathfrak{E}} \end{aligned} \tag{41}$$

where we will call $\mathbf{t}_i^{MPRR}$ the MPRR /RST /VAT /GST component of consumption tax t_i and $t_i^{\mathfrak{E}} = t_i - \mathbf{t}_i^{MPRR}$ is called its externality excise component. For reasons to be specified in Section 4.3, we will call $t_{\mathbb{o}}$ and $t_{\mathbb{R}}$ as oil and road excises.

The following corollary to Theorem 8 establishes the links between various externality excises at the second-best optimum $\bar{\Upsilon}$.

Corollary 9 *Suppose at the second-best optimum $\bar{\Upsilon}$, we have $\bar{\eta} \in \Omega_\eta$ and $\bar{t} \in \Omega_t(\bar{\eta})$. Then the second-best optimal externality excises on various goods satisfy*

$$\overline{t}_i^{\mathfrak{E}} = 0 \quad \forall\, i \in \widehat{\mathcal{N}} \cup \{\mathbb{e}\} \tag{42}$$

$$\overline{t}_i^{\mathfrak{E}} + \alpha_i \bar{t}_{\mathbb{o}} + \beta_i \bar{t}_{\mathbb{R}} + \alpha_i \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}} = \beta_i \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} + \alpha_i \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b} \tag{43}$$

$$\overline{t}_{\mathbb{k}}^{\mathfrak{E}} + \alpha_{\mathbb{k}} \bar{t}_{\mathbb{o}} + \alpha_{\mathbb{k}} \phi_{\mathbb{u}} \bar{\eta}_{\mathbb{u}} = \alpha_i \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}} \tag{44}$$

$$\overline{t}_{\mathbb{c}}^{\mathfrak{E}} + \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} = \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \tag{45}$$

$$\overline{t}_{\mathbb{w}_i}^{\mathfrak{E}} + \phi_{\mathbb{c}\mathbb{w}_1} \bar{\eta}_{\mathbb{c}\mathbb{w}_i} + \phi_{\mathbb{c}\mathbb{w}_i} \phi_{\mathbb{v}} \bar{\eta}_{\mathbb{v}} = \frac{\bar{\gamma}}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{c}\mathbb{w}_i} \quad \forall\, i = 1, \ldots, w \tag{46}$$

$$\overline{t}_{\mathbb{l}}^{\mathfrak{E}} + \phi_{\mathbb{t}} \bar{\eta}_{\mathbb{t}\mathbb{l}} = \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{l}} \tag{47}$$

$$\overline{t}_{\mathbb{t}}^{\mathfrak{E}} = \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{t}} \tag{48}$$

On the other hand, DM's characterisation of the MPRR follows in conventional way from

([34](#)) for all commodities in the index set $\mathcal{N}$ except $\mathbb{o}$ and $\mathbb{R}$:[46]

$$\frac{\sum_h \sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \frac{\partial x_j^h}{\partial \delta_i}\mathbf{t}_i^{MPRR}}{\sum_h x_j^h} = -1+\sum_h \frac{\mu^h\lambda^h}{\kappa}\frac{x_j^h}{\sum_h x_j^h}+\sum_h \sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \frac{\partial x_i^h}{\partial m}\mathbf{t}_i^{MPRR}\frac{x_j^h}{\sum_h x_j^h} - \sum_h \epsilon_{jm}^h \frac{\sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}}\delta_i x_i^h}{m+m^h}\frac{x_j^h}{\sum_h x_j^h} \tag{49}$$

where for all $h = 1,\ldots,H$, ϵ_{jm}^h is the income elasticity of demand for the j^{th} good in the index set $\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}$.

4.3 A discussion of Theorem 8, Corollary 9, and the MPRR.

Theorem 8 demonstrates the additive nature of the VAT/GST and the externality excise for most goods at a second-best optimum. The consumption tax on a commodity is a sum of a MPRR component that takes into account equity-efficiency considerations and a Pigouvian tax component that addresses the externality correction. The latter component is zero if the commodity is not associated with the generation of externalities, *e.g.,* it is zero for commodities in $\widehat{\mathcal{N}}$.

In contrast to the above second-best optimal consumption tax structure, real life policies often employ consumption tax rates that are multiplicative in the MPRR and the externality-excise components. For example, in many countries such as UK and India, cigarette excise consists of both a unit and an ad-valorem component, and the VAT/GST is levied as a certain percentage of the pack price inclusive of the externality excises. This procedure is also adopted for real life taxation of alcohol and motor fuels.

4.3.1 The MPRR component of the consumption tax.

For any good $j\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}$, the component $\mathbf{t}_j^{MPRR}$ of the consumption tax rate t_j satisfies the well-known MPRR in ([49](#)), namely, the percentage decline in the demand for good j due to the tax system $\mathbf{t}_j^{MPRR}$ is more (i) the more is the share of people with lower welfare weights (measured in terms of the social marginal utilities of income $\mu^h\lambda^h$) in the total consumption of the good and (ii) the more is the consumption of the good concentrated among people whose tax payments are less responsive to commodity-tax induced changes in real income. Point (i)

[46] The common formulation of the MPRR is in terms of percentage reductions in compensated demands for different commodities (see *e.g.,* Myles (1995) and Atkinson and Stiglitz (1976, 1980)):

$$\frac{\sum_h \sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \frac{\partial \mathfrak{H}_j^h}{\partial \delta_i}\mathbf{t}_i^{MPRR}}{\sum_h x_j^h} = -1+\sum_h \frac{\mu^h\lambda^h}{\kappa}\frac{x_j^h}{\sum_h x_j^h}+\sum_h \sum_{i\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \frac{\partial x_i^h}{\partial m}\mathbf{t}_i^{MPRR}\frac{x_j^h}{\sum_h x_j^h}$$

where $\mathfrak{H}(\delta,u)_j^h$ denotes the j^{th} compensated demand for consumer h. The formulation of the MPRR given here in terms of percentage reductions in uncompensated demands for different commodities is obtained from DM (b), p. 268, eqn. 77.

addresses equity considerations, while point (ii) addresses efficiency considerations.[47]

Note also that the theorem does not specify any MPRR components of consumption taxes for oil and road services, $t_{\mathbb{o}}$ and $t_{\mathbb{R}}$. (See also Figure 2 for the points at which consumers are taxed for vehicular services used.) No separate MPRR tax needs to be imposed on these goods since the demands for these goods are completely derived from the demands for goods $\mathbb{s}$, $\mathbb{m}$, $\mathbb{b}$, and $\mathbb{k}$. Discouraging consumption of these latter goods in line with the MPRR automatically ensures discouragement of goods $\mathbb{o}$ and $\mathbb{R}$ in line with the MPRR. To see this, consider for example, the case of oil. Taking equations corresponding to $j = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$ in (34) and summing up these equations after multiplying both sides of the j^{th} equation by α_j, we obtain the following by exploiting the Slutsky equation and symmetry of the Slutsky matrix

$$\frac{\sum_h \sum_{i \in \mathbb{N}\setminus\{\mathbb{o},\mathbb{R}\}} \frac{\partial \mathfrak{H}^h_{\mathbb{o}}}{\partial \delta_i} \mathbf{t}_i^{MPRR}}{\sum_h x^h_{\mathbb{o}}} \;=\; -1 + \sum_h \frac{\mu^h \lambda^h}{\bar{\kappa}} \frac{x^h_{\mathbb{o}}}{\sum_h x^h_{\mathbb{o}}} + \sum_h \sum_{i \neq \mathbb{o}} \frac{\partial x^h_i}{\partial m} \mathbf{t}_i^{MPRR} \frac{x^h_{\mathbb{o}}}{\sum_h x^h_{\mathbb{o}}}.$$

Thus, MPRR component of the optimal tax structure $\mathbf{t}^{MPRR}$ is such that the compensated demand for oil is discouraged keeping in mind equity and efficiency considerations. Hence, $\bar{t}_{\mathbb{o}}$ and $\bar{t}_{\mathbb{R}}$ in Theorem 8 and Corollary 9 can be interpreted purely as externality excises on oil and road services.

The same argument can also be extended to the case of electricity. Electricity helps in the functioning of many electrical and electronic devices, from which work is extracted in production and consumption. If we modify our model to capture this feature of electricity then, as in the case of oil and road, consumption demand for electricity will be indirectly derived from the demand for all electrical equipment and electronic gadgets used by consumers. Similarly, the use of electricity as an input by firms will be complementary to their use of electrical and electronic devices as inputs. Suppose $\mathbb{E}_1, \ldots, \mathbb{E}_e$ is a list of all devices requiring electricity for their functioning and suppose a unit service from the i^{th} such device requires $\phi_{\mathbb{E}_i}$ amount of electricity. In this case, at the second-best $\bar{\Upsilon}$, the consumption and intermediate input tax rates on these devices and electricity can be shown to be related in the following manner:

$$\begin{aligned} \bar{t}_{\mathbb{E}_i} + \bar{t}_{\mathbb{e}} \phi_{\mathbb{E}_i} - \bar{\eta}_{\mathbb{E}_i} &= \bar{\mathbf{t}}^{MPRR}_{\mathbb{E}_i} \quad \forall\, i = 1, \ldots, e, \qquad \eta_{\mathbb{E}_i} + \phi_{\mathbb{E}_i} \eta_{\mathbb{r}} \;=\; 0 \\ \eta_{\mathbb{E}_i} + \phi_{\mathbb{E}_i} \eta_{\mathbb{n}} + \phi_{\mathbb{E}_i} \phi_{\mathbb{cn}} \eta_{\mathbb{cn}} + \phi_{\mathbb{E}_i} \phi_{\mathbb{cn}} \phi_{\mathbb{v}} \eta_{\mathbb{v}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{c}} \phi_{\mathbb{cn}} \phi_{\mathbb{E}_i} \;\forall\, i = 1, \ldots, e. \end{aligned}$$

The total consumption of electricity by consumer h is the sum of electricity required to run the electrical and electronic devices he demands, *i.e.*, $x^h_e = \sum_{i=1}^{e} \phi_{\mathbb{E}_i} x^h_{\mathbb{E}_i}$. The vector $\bar{\mathbf{t}}^{MPRR}$ will now includes the MPRR tax rates on devices that run on electricity. Since $\bar{\mathbf{t}}^{MPRR}$ discourages the consumption of devices that run on electricity along the lines of the MPRR, it also discourages the consumption of electricity along the principles of the MPRR. Hence, this argument implies that, at a second-best optimum, there is no need for a GST or VAT or RST on electricity, if

[47]Note that, unlike in MP, this component of the consumption tax is independent of the externality component.

GST is implemented on devices that run using electricity.

4.3.2 The externality excise component of the consumption tax.

Corollary 9 describes the linkages between the externality excise components of the consumption taxes and the intermediate input tax rates. For example, the conditions in (43) show that the monetised value of the SMC (from congestion and carbon externalities) of motoring service from car of size $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$ is given by $\beta_i \frac{\bar{\gamma}_C}{\bar{\kappa}} \psi_{\mathbb{R}} + \alpha_i \frac{\bar{\gamma}_Y}{\bar{\kappa}} \psi_{\mathbb{o}}$. This is also the increase in social cost from use of α_i amount of oil and β_i amount of road service used in driving car of type i, as well as the increase in social cost due to use of $\alpha_i \phi_{\mathbb{u}}$ amount of unrefined oil for producing the refined oil needed for deriving a unit service from car of type i. Thus, given the intermediate input tax on crude oil $\bar{\eta}_{\mathbb{u}}$, the externality excise rates for consumers $\bar{t}_i^{\mathfrak{E}}$ and $t_{\mathbb{o}}$ should be chosen such that the sum of the Pigouvian tax revenues from taxing consumers for a unit service consumed from car of type i, for the road service and refined oil needed for deriving a unit service from car of type i, as well as the revenue from taxing the sale of unrefined oil needed to produce the refined oil needed for deriving a unit service from car of type i should be equal to the monetised value of the SMC. Conditions in (43) show that use of a few or all of these tax rates is permitted at the second-best optimum.

Condition (47) shows that consumption externality from cigarettes can be controlled by either taxing cigarettes or taxing or the use of tobacco for making cigarettes or both.

Note also that (43) and (44) indicate several degrees of freedom in choosing tax rates $\bar{t}_i^{\mathfrak{E}}$ for $i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}$; $\bar{t}_{\mathbb{o}}$; and $\bar{t}_{\mathbb{R}}$. As in the case of Theorem 6, consumption taxation of oil and road service is not enough to combat congestion and carbon externalities of motoring generated by consumers, when there are significant variations in the types of the vehicles and non-motoring uses of oil.

Further, Theorems 6 and 8 indicate that externality taxation of many commodities need not be equal in consumption and production, *e.g.*, $\bar{t}_{\mathbb{o}}, \bar{t}_{\mathbb{R}}, \bar{t}_{\mathbb{w}}^{\mathfrak{E}}$ could be chosen to be different from $\bar{\eta}_{\mathbb{o}}, \bar{\eta}_{\mathbb{R}}, \bar{\eta}_{\mathbb{w}}^{\mathfrak{E}}$, respectively.

5 Interpretations of Lagrange multipliers γ_Y, γ_X, and γ_C.

From the first-order conditions of the second-best welfare maximisation with respect to z_Y, z_C, and z_X, we obtain the following characterisations of the Lagrange multipliers of the Nash equilibrium constraints on the externalities:

$$
\begin{aligned}
\bar{\gamma}_i &= -\sum_h \mu_h u^h_{z_i} - \left[\sum_{j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\}} \bar{\kappa} \bar{\mathbf{t}}_j^{MPRR} \frac{\partial x_j}{\partial z_i} - \left(-\bar{\kappa} \sum_g \bar{\tau}_g \frac{\partial \Pi^g}{\partial z_i} \right) \right] \qquad \forall\ i = C, Y \\
\bar{\gamma}_X &= -\sum_h \mu_h u^h_{z_X} - \sum_{j \in \mathcal{N} \setminus \{\mathbb{o}, \mathbb{R}\}} \bar{\kappa} \bar{t}_j^{MPRR} \frac{\partial x_j}{\partial z_X} \qquad (50)
\end{aligned}
$$

For $\iota = X, C, Y$, the Lagrange multiplier $\bar{\gamma}_\iota$ measures the impact on social welfare of a positive marginal deviation from the Nash equilibrium constraint that externality ι is required to satisfy at the optimum of the welfare maximisation problem (16). It can also be interpreted as the net social marginal cost of externality ι. This is because from (50) it follows that $\bar{\gamma}_\iota$ is the social marginal disutility to consumers from the externality (given by $-\sum_h \mu_h u^h_{z_\iota} > 0$) minus the social marginal benefits from the fiscal effects generated by the externality.[48] The latter is the social value of the increase in the commodity tax revenue to the government (given by $\sum_{j\in\mathcal{N}\setminus\{\mathbb{o},\mathbb{R}\}} \bar{\kappa}\bar{\mathbf{t}}_j^{MPRR}\frac{\partial x_j}{\partial z_i}$) net of the social value of the decrease in profit tax revenue from a unit increase in the level of the detrimental externality (given by $-\bar{\kappa}\sum_g \bar{\tau}_g \frac{\partial \Pi^g}{\partial z_i} > 0$).[49]

Note that the fiscal effects generated by externalities imply that the value of the Lagrange multipliers $\bar{\gamma}_Y$, $\bar{\gamma}_C$, and $\bar{\gamma}_X$ cannot be signed unambiguously. Significant fiscal effects resulting in an increase in the commodity tax revenue can, in principle, offset the detrimental effects of externality generation – namely, the increase in disutility to consumers and reductions in profits of firms. However, if fiscal effects are small, then these multipliers will take positive values.

6 A system of value added taxes with input tax credits.

The optimal intermediate input taxes and the consumption taxes identified in Theorems 6 and 8 can be implemented as a system of VAT/GST with input tax credits, where (i) both producers and consumers buy goods in the index set $\mathcal{N}$ (either as intermediate inputs or for consumption) at the retail prices q and (ii) producers receive some tax credits for the taxes paid during the purchase of inputs.

Definition. Suppose $\langle \delta, \rho_O, \rho_I, m, \tau_1, \ldots, \tau_G, z_C, z_X, z_Y\rangle \in \mathbf{R}^{N+\mathfrak{N}_\rho+G+2}$ is a tax equilibrium. Let $p \in \mathcal{P}(\rho)$ and $q \in \mathcal{Q}(\delta)$. Suppose $\eta \in \mathbf{R}^{\mathfrak{N}_O+2}$ and $t \in \mathbf{R}^N$ are, respectively, the vectors of intermediate input and consumption tax rates associated with p and q. Then $\langle \eta, t\rangle \in \mathbf{R}^{\mathfrak{N}_O+N+2}$ is implemented as a system of value-added taxes with input tax credits if producers face price vector $\Gamma = \langle \Gamma_O, \Gamma_I\rangle \in \mathbf{R}_+^{\mathfrak{N}_O} \times \mathbf{R}_+^{\mathfrak{N}_I}$ where

$$\begin{array}{rclrclrcl} \Gamma_{Ii} & = & q_i \ \forall\ i \in \mathcal{N}\setminus\{\mathbb{t}, \mathbb{c}\}; & \Gamma_{I\mathbb{t}i} & = & q_{\mathbb{t}} \text{ for } i = \mathbb{l}, \mathbb{a}; & & & \\ \Gamma_{I\mathbb{c}i} & = & q_{\mathbb{c}} \text{ for } i = \mathbb{n}, \mathbb{w}_1, \ldots, \mathbb{w}_w; & \Gamma_{Ii} & = & p_{Ii} \text{ for } i = \mathbb{u}, \mathbb{v}; & \Gamma_O & = & p_O \end{array}$$

and receive the following credits for taxes paid on goods in index set $\mathcal{N}$ when they are purchased

[48] Recall that $\bar{\kappa}$ is the social marginal value of money/numeraire.

[49] Recall, that the profits of firms are adversely affected by the carbon and congestion externalities. Such fiscal effects of externality generation or public good production at a second-best have also been discussed by Atkinson and Stern (1974) and Bovenberg and Ploeg (1994). Fiscal effects arise because of the substitutability or complementarity relations between the externality and various consumption goods. For example, the adverse health effect of cigarette smoking can increase medical expenditures, which may be subject to commodity taxes.

as inputs:

$$
\begin{aligned}
\Delta_i &= q_i - p_{Ii} \quad \forall\, i \in \mathcal{N} \setminus \{\mathbb{c}, \mathbb{t}\}; \\
\Delta_{\mathbb{t}i} &= q_{\mathbb{t}} - p_{I\mathbb{t}i} \quad \forall\, i = \mathbb{l}, \mathbb{a}; \qquad\qquad \Delta_{\mathbb{c}j} = q_{\mathbb{c}} - p_{I\mathbb{c}i} \;\forall\, i = \mathbb{n}, \mathbb{w}_1, \ldots, \mathbb{w}_w.
\end{aligned}
$$

In particular, consider the second-best optimum $\bar{\Upsilon}$. Suppose $\bar{p} \in \mathcal{P}(\bar{\rho})$, $\bar{q} \in \mathcal{Q}(\delta)$, and $\bar{\eta} \in \Omega_\eta$ and $\bar{t} \in \Omega_t(\bar{\eta})$ are the associated vectors of intermediate-input and consumption tax rates, respectively. Then, (41) implies that for all $i \in \mathcal{N} \setminus \{\mathbb{e}\}$, we have

$$
\begin{aligned}
\bar{\Delta}_i &= \bar{q}_i - \bar{p}_{Ii} = (\bar{q}_i - \bar{p}_{Oi}) + (\bar{p}_{Oi} - \bar{p}_{Ii}) \\
&= \bar{t}_i - \bar{\eta}_i = \bar{\mathfrak{t}}_i^{MPRR} + \left[\bar{t}^{\mathfrak{E}} - \bar{\eta}_i\right], \qquad (51)
\end{aligned}
$$

while for electricity, $\bar{\Delta}_i = \bar{t}_{\mathbb{e}} = \bar{\mathfrak{t}}_{\mathbb{e}}^{MPRR}$. Thus, the input tax credit to be received by a producer on his purchase of any input (other than electricity) at consumer price, has two components – the MPRR component of the consumption tax on that good and the difference between the externality excise paid by the consumer and the externality excise that is due from the producer. Based on this, Theorem 10 below, whose conclusions follow in a straightforward way, provides the rates of input tax credits due to the producers on various purchases of inputs.

Theorem 10 *At the second-best optimum $\bar{\Upsilon}$, let $\bar{p} \in \mathcal{P}(\bar{\rho})$ and $\bar{q} \in \mathcal{Q}(\delta)$. Suppose $\bar{\eta} \in \Omega_\eta$ and $\bar{t} \in \Omega_t(\bar{\eta})$ are the associated vectors of intermediate input taxes and consumption tax rates, respectively. Suppose $\langle \bar{\eta}, \bar{t} \rangle \in \mathbf{R}^{\mathfrak{N}_O + N + 2}$ is implemented as a system of value-added taxes with input tax credits. Then producers receive the following rates of input-tax credits:*

$$
\begin{aligned}
\bar{\Delta}_i &= \bar{\mathfrak{t}}_i^{MPRR} \quad \forall\, i \in \widehat{\mathcal{N}} \cup \{\mathbb{e}, \mathbb{w}_1, \ldots, \mathbb{w}_w\} \\
\bar{\Delta}_i &= \bar{\mathfrak{t}}_i^{MPRR} - \alpha_i \left[\bar{t}_{\mathbb{o}} - \bar{\eta}_{\mathbb{o}}\right] - \beta_i \left[\bar{t}_{\mathbb{R}} - \bar{\eta}_{\mathbb{R}}\right] \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b} \\
\bar{\Delta}_{\mathbb{k}} &= \bar{\mathfrak{t}}_{\mathbb{k}}^{MPRR} - \alpha_{\mathbb{k}} \left[\bar{t}_{\mathbb{o}} - \bar{\eta}_{\mathbb{o}}\right], \qquad \bar{\Delta}_i = \bar{t}_i^{\mathfrak{E}} - \bar{\eta}_i = \bar{t}_i - \bar{\eta}_i \quad \forall\, i = \mathbb{o}, \mathbb{R}, \\
\bar{\Delta}_{\mathbb{t}\mathbb{a}} &= \bar{\mathfrak{t}}_{\mathbb{t}}^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{t}}, \qquad \bar{\Delta}_{\mathbb{t}\mathbb{l}} = \bar{\mathfrak{t}}_{\mathbb{t}}^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{t}} - \bar{\eta}_{\mathbb{t}\mathbb{l}} = \bar{\mathfrak{t}}_{\mathbb{t}}^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{t}} + \frac{\bar{\eta}_{\mathbb{l}}}{\phi_{\mathbb{t}}} \\
\bar{\Delta}_{\mathbb{l}} &= \bar{\mathfrak{t}}_{\mathbb{l}}^{MPRR} + \frac{\bar{\gamma}_X}{\bar{\kappa}} \psi_{\mathbb{l}}, \qquad \bar{\Delta}_i = \bar{\mathfrak{t}}_{\mathbb{c}}^{MPRR} + \frac{\bar{\eta}_i}{\phi_{\mathbb{c}i}} \qquad \forall\, i = \mathbb{n}, \mathbb{w}_1, \ldots, \mathbb{w}_w
\end{aligned}
$$

If $\bar{\mathfrak{t}}_i = \bar{\eta}_i$ for $i \in \{\mathbb{o}, \mathbb{R}\}$, then we also have

$$
\bar{\Delta}_i = \bar{\mathfrak{t}}_i^{MPRR} \quad \forall\, i \in \{\mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}\}
$$

Thus, at a second-best optimum, producers will not receive full credit on the taxes paid on goods causing production externalities; rather credit is received on (i) the MPRR components of the taxes and (ii) the difference between the externality excises paid by the consumer and that due from the producers. Contrast this with the input-tax credit received in some real-life systems of VAT, where VAT and externality excises are applied in a multiplicative way, as for example in the case of taxation of motoring fuel in UK and India, which imply that the carbon-externality generating producers will be reimbursed on both the VAT and externality excise components.

In the case of goods generating consumption externalities (such as goods $\mathbb{l}$ and $\mathbb{t}$) , because producers are not the generators, Theorem 10 shows that the optimal tax credits to producers include both the MPRR and the externality component of the consumption tax.

7 Extensions.

The model studied can be extended along many dimensions. In this section, it is extended to incorporate and distinguish between two types of subsidy policies. The first is Pigouvian in nature that is intended to encourage production of a good or service (such as carbon capture and sequestration (CCS)) that generates a positive externality, while the second encourages production of a good beyond its short-run optimal level, with an aim of promoting long-run sustainable development. Annual budgetary and fiscal plans of the government are usually based on the current economic and technological environments, and are intended for meeting not only its current revenue, redistribution, and externality correction objectives, but also for realising its vision for the future course of growth and development of the economy. If the latter is one of sustainable development, then some of its current choices of policies may appear suboptimal from today's point of view, but may provide incentives to producers to gradually shift to adopting practices that promote its long-run objectives. Such policies may include quantity targets for renewable energy production or use of cleaner fuels. Subsidies can be used as policy instruments for realising these quantity targets through the market mechanism.

7.1 A cleaner motoring fuel with a quantity target.

The set of commodity indexes $\mathcal{N}$ is extended to include an alternative motoring fuel denoted by $\mathbb{g}$ (*e.g.,* CNG) that is less emission intensive. To keep the presentation simple and focused, without loss of generality, we assume that the smaller-sized vehicle can be operated by using either of the two motoring fuels $\mathbb{o}$ and $\mathbb{g}$. Services consumed by consumer h of small sized vehicles by using oil and CNG are denoted by $x^h_{\mathbb{so}}$ and $x^h_{\mathbb{sg}}$, respectively. If these motoring services are perfectly substitutable in consumption with consumer utility depending on total services from small-sized vehicles $x^h_{\mathbb{so}} + x^h_{\mathbb{sg}}$ then, with unequal market prices of oil and CNG, consumers will choose to consume only the cheaper of the two, with consumption of the other being zero. To allow the possibility of non-zero consumption demands for both fuels, we assume that consumers also derive some warm glow (feel good) utility from consuming the more environment-friendly fuel CNG. Thus, consumer utility is a function of both $x^h_{\mathbb{so}} + x^h_{\mathbb{sg}}$ and $x^h_{\mathbb{sg}}$.

$$u^h = u^h\left(x^h_{\mathcal{N}\setminus\{\mathbb{o},\mathbb{g},\mathbb{R},\mathbb{so},\mathbb{sg}\}}, x^h_{\mathbb{so}} + x^h_{\mathbb{sg}}, x^h_{\mathbb{sg}}, z_C, z_X, z_Y\right).$$

The demand for oil, CNG, and road services is derived from the demand for services from vehicles and from demand for other non-motoring services that use fuel:

$$x_{\mathbb{o}}^h = \alpha_{\mathbb{so}} x_{\mathbb{so}}^h + \sum_{i=\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i x_i^h, \quad x_{\mathbb{g}}^h = \alpha_{\mathbb{sg}} x_{\mathbb{sg}}^h, \quad \text{and} \quad x_{\mathbb{R}}^h = \sum_{i=\mathbb{s},\mathbb{m},\mathbb{b}} \beta_i x_i^h,$$

where $\alpha_{\mathbb{so}}$ and $\alpha_{\mathbb{sg}}$ denote the amounts of fuels $\mathbb{o}$ and $\mathbb{g}$ required for a unit service from vehicle of type $\mathbb{s}$. Similarly, technological feasibility requires that for any firm f, we have $y_{I\mathbb{o}}^f = \alpha_{\mathbb{so}} y_{I\mathbb{so}}^f + \sum_{i=\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i y_{Ii}^f$ and $y_{I\mathbb{g}}^f = \alpha_{\mathbb{sg}} y_{I\mathbb{sg}}^f$, where $y_{I\mathbb{s}i}^f$ denotes the amount of services taken by firm f from small vehicles driven by fuel of type $i = \mathbb{o}, \mathbb{g}$. Suppose $\psi_{\mathbb{g}}$ denotes the CO_2-emission intensity of CNG. With this extension to the model, it can be shown that the second-best optimal intermediate input and consumption taxation of services from small-sized vehicles satisfy the following modifications of conditions (26) and (36) of Theorems 6 and 8, respectively:

$$\begin{aligned}
\eta_{\mathbb{so}} + \alpha_{\mathbb{so}}\eta_{\mathbb{o}} + \beta_{\mathbb{s}}\eta_{\mathbb{R}} + \alpha_{\mathbb{so}}\phi_{\mathbb{u}}\eta_{\mathbb{u}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{o}}\alpha_{\mathbb{so}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}} \\
\eta_{\mathbb{sg}} + \alpha_{\mathbb{sg}}\eta_{\mathbb{g}} + \beta_{\mathbb{s}}\eta_{\mathbb{R}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}}\alpha_{\mathbb{sg}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}} \\
t_{\mathbb{so}} + \beta_{\mathbb{s}}\mathbb{t}_{\mathbb{R}} + \alpha_{\mathbb{so}}t_{\mathbb{o}} + \alpha_{\mathbb{so}}\phi_{\mathbb{u}}\bar{\eta}_{\mathbb{u}} &= \bar{\mathbf{t}}_{\mathbb{so}}^{MPRR} + \beta_{\mathbb{s}}\frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}} + \alpha_{\mathbb{so}}\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{o}} \\
t_{\mathbb{sg}} + \beta_{\mathbb{s}}\mathbb{t}_{\mathbb{R}} + \alpha_{\mathbb{sg}}t_{\mathbb{g}} &= \bar{\mathbf{t}}_{\mathbb{sg}}^{MPRR} + \beta_{\mathbb{s}}\frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}} + \alpha_{\mathbb{sg}}\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}} \qquad (52)
\end{aligned}$$

Given the warm glow effect from usage of CNG, the MPRR components of the consumption taxes on services of small cars driven by oil and CNG will not be the same, *i.e.*, $\bar{\mathbf{t}}_{\mathbb{so}}^{MPRR} \neq \bar{\mathbf{t}}_{\mathbb{sg}}^{MPRR}$. That is, the extent of discouragement based on equity and efficiency considerations by the consumption tax system of consumption of vehicular services will vary depending on the motoring fuel used. Further, given differences in the amount of emission generated by oil and CNG per unit usage of services from small cars (*i.e.*, given that $\alpha_{\mathbb{so}}\psi_{\mathbb{o}} \neq \alpha_{\mathbb{sg}}\psi_{\mathbb{g}}$), the social marginal costs of using oil and CNG to run small cars (given by $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{o}}\alpha_{\mathbb{so}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}}$ and $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}}\alpha_{\mathbb{sg}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}}$) will differ. In particular, since CNG is a cleaner fuel, one expects $\alpha_{\mathbb{so}}\psi_{\mathbb{o}} > \alpha_{\mathbb{sg}}\psi_{\mathbb{g}}$, so that the social marginal cost will be lower when CNG is used to drive small cars. Hence, the optimal intermediate input and consumption tax revenues from using CNG or oil per unit service from small cars (given by the left sides of conditions in (52)) will differ. In particular, lower taxation of CNG is recommended if $\bar{\mathbf{t}}_{\mathbb{so}}^{MPRR} > \bar{\mathbf{t}}_{\mathbb{sg}}^{MPRR}$ in addition to $\alpha_{\mathbb{so}}\psi_{\mathbb{o}} > \alpha_{\mathbb{sg}}\psi_{\mathbb{g}}$.

Suppose the second-best optimal level of CNG usage at the second-best optimum is $\bar{y}_{O\mathbb{g}}$. If a quantity constraint equal to $\tilde{y}_{O\mathbb{g}}$ is imposed on CNG to promote its usage beyond this level, then the required tax equilibrium should also satisfy

$$\sum_h x_{\mathbb{g}}^h + \sum_g y_{I\mathbb{g}}^g \geq \tilde{y}_{O\mathbb{g}},$$

where $\tilde{y}_{O\mathbb{g}} > \bar{y}_{O\mathbb{g}}$. If we now solve the welfare maximisation problem (16) with this additional constraint then, at the quantity-constrained optimum, this constraint is binding (so that the value of the Lagrange multiplier of this constraint evaluated at the optimum and denoted by $\bar{\omega}_{\mathbb{g}}$ is positive), and the optimal intermediate input and consumption taxation of services from small-sized vehicles when CNG is used are modified to

$$\begin{aligned} \eta_{\mathbb{sg}} + \alpha_{\mathbb{sg}}\eta_{\mathbb{g}} + \beta_{\mathbb{s}}\eta_{\mathbb{R}} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}}\alpha_{\mathbb{sg}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}} - \alpha_{\mathbb{sg}}\frac{\bar{\omega}_{\mathbb{g}}}{\bar{\kappa}} \\ t_{\mathbb{sg}} + \alpha_{\mathbb{sg}}t_{\mathbb{g}} + \beta_{\mathbb{s}}t_{\mathbb{R}} &= \bar{\mathbf{t}}^{MPRR}_{\mathbb{sg}} + \beta_{\mathbb{s}}\frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}} + \alpha_{\mathbb{sg}}\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}} - \alpha_{\mathbb{sg}}\frac{\bar{\omega}_{\mathbb{g}}}{\bar{\kappa}} \end{aligned} \tag{53}$$

Thus, in contrast to the analogous condition in (52), when there is a quantity constraint on a cleaner motoring fuel such as CNG, then the intermediate-input tax revenue collected (given by the left-side of the first equation in (53)) is smaller than the monetised value of the SMC of the externality generated when a unit of service is derived from the small-sized vehicle run on CNG (where the SMC is given by $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{g}}\alpha_{\mathbb{sg}} + \frac{\bar{\gamma}_C}{\bar{\kappa}}\psi_{\mathbb{R}}\beta_{\mathbb{s}}$). This provides a price incentive to the firms to use more CNG. Similarly, to incentivise consumers to use CNG, the consumption tax revenue collected when a unit of service is derived from small-sized vehicles run on CNG is smaller than at the quantity-unconstrained optimum characterised in (52). Thus, implementing the quantity constraint on CNG requires a subsidy equal to $\alpha_{\mathbb{sg}}\frac{\bar{\omega}_{\mathbb{g}}}{\bar{\kappa}}$, which lowers the market price of CNG faced by buyers.

7.2 A quantity constraint on renewable energy generation.

Suppose $\bar{y}_{O\mathbb{r}}$ is the second-best optimal level of renewable energy obtained by solving the welfare maximisation problem (16) . Note, it is possible that, at an optimum of problem (16), zero amount of renewable energy is produced. This will be true if the cost of generation of energy from renewables is too high relative to the market price $p_{O\mathbb{r}}$ received by its producers. In that case, profit maximisation by these firms may lead to zero production.

If a quantity constraint equal to at least $\tilde{y}_{O\mathbb{g}}$ amount of renewable energy production

$$\sum_g y^g_{O\mathbb{r}} \geq \bar{y}_{O\mathbb{r}}$$

is now imposed and the welfare maximisation problem (16) is solved with this additional constraint, then the constraint is binding at the optimum only if $\tilde{y}_{O\mathbb{r}} > \bar{y}_{O\mathbb{r}}$. It can be shown that the optimal intermediate input tax on renewables is now given by

$$\bar{\eta}_{\mathbb{r}} = -\frac{\bar{\omega}_{\mathbb{r}}}{\bar{\kappa}} < 0$$

where $\bar{\omega}_{\mathbb{r}} > 0$ is the Lagrange multiplier of the quantity constraint on renewable energy, evaluated at the optimum. Recall that renewable energy is already subsidised at the second-best optimum $\bar{\Upsilon}$ of the original problem (16) without the quantity constraint. This is because, as seen in Theorem 6, in this case, the optimal intermediate input tax on renewable electricity is zero as it does not contribute to the carbon externality, while it is equal to the marginal cost of the carbon externality in the case of thermal electricity generation. Now, in addition to this implicit subsidy, a further subsidy equal to $\frac{\bar{\omega}_{\mathbb{r}}}{\bar{\kappa}} > 0$ has to be given to the producers to incentivise them to generate the targeted amount of renewable electricity. Thus, the price of renewable energy received by the producers is now

$$p_{O\mathbb{r}} = p_{I\mathbb{e}} + \frac{\bar{\omega}_{\mathbb{r}}}{\bar{\kappa}},$$

where, recall, $p_{I\mathbb{e}}$ is the common price of electricity (immaterial of the source – renewable or non-renewable) paid by electricity-purchasing firms.

7.3 Alternative sources of thermal electricity generation and allowing for carbon sequestration techniques.

The model can be extended to allow for technological differences in thermal electricity generation based on different varieties of fossil-fuels (here coal) used as well as carbon capture and sequestration techniques. Without loss of generality, assume that there are two varieties of thermal power plants indexed by $\mathbb{n}_1$ and $\mathbb{n}_2$, each using a different variety of coal ($\mathbb{c}_1$ and $\mathbb{c}_2$) to produce electricity. The outputs of the mining sectors used as inputs into the production of these two types of coal are $\mathbb{v}_1$ and $\mathbb{v}_2$, respectively. The fossil-fuel externality can also be reduced by carbon capture and sequestration (CCS) efforts of firms indexed by $\mathbb{S}$. This modifies the the carbon externality generation equation to

$$\begin{aligned} z_Y \;&=\; \psi_{\mathbb{o}}\left[\sum_{i=\mathbb{s},\mathbb{m},\mathbb{b},\mathbb{k}} \alpha_i \left(\sum_h x_i^h + \sum_g y_{Ii}^g\right)\right] + \sum_{i=1,2} \psi_{\mathbb{c}_i}\left[\sum_h x_{\mathbb{c}_i}^h + \sum_g y_{I\mathbb{c}_i\mathbb{a}}^g + \phi_{\mathbb{c}_i}\sum_g y_{O\mathbb{n}_i}^g\right] \\ &\quad - \psi_{\mathbb{S}}\sum_g y_{I\mathbb{S}}^g \end{aligned} \tag{54}$$

where $\psi_{\mathbb{c}i}$ is the CO_2 emission intensity of carbon of type $i = 1, 2$ and $\psi_{\mathbb{S}}$ is the amount of mitigation of carbon emission done by one unit of CCS effort. The important point to note is that CCS generates a positive externality for the society. This is because, if CCS level increases, then (54) shows that the the carbon externality level z_Y decreases. Hence, since utilities of consumers and profits of firms are assumed to be adversely affected by fossil-fuel externality z_Y, an increase in the total CCS effort by firms implies increases in the utilities of consumers and the profits of firms.

We can distinguish between firms that are producing and those that are demanding CCS services (it is possible that a firm both demands and produces CCS). The tax equilibrium is modified to include the following market clearing condition for CCS:

$$\sum_g y_{O\mathbb{S}}^g - \sum_g y_{I\mathbb{S}}^g = 0$$

If the output and input prices of sequestration are $p_{O\mathbb{S}}$ and $p_{I\mathbb{S}}$, respectively, then the intermediate input tax on sequestration is $\eta_{\mathbb{S}} = p_{I\mathbb{S}} - p_{O\mathbb{S}}$. With these extensions to the model, the set of conditions (24) of Theorem 6 are modified as follows:

$$\begin{aligned} \bar{\eta}_{\mathbb{m}_i} + \phi_{\mathbb{c}_i}\bar{\eta}_{\mathbb{c}_i\mathbb{m}} + \phi_{\mathbb{v}_i}\phi_{\mathbb{c}_i}\bar{\eta}_{\mathbb{v}_i} &= \frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{c}_i}\phi_{\mathbb{c}_i} \qquad \forall\ i = 1, 2 \\ \bar{\eta}_{\mathbb{S}} &= -\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{S}} < 0 \end{aligned} \tag{55}$$

Thus, the above conditions show that, at a second-best optimum, the Pigouvian tax revenue collected when a unit of thermal electricity is produced by using coal of type $i = 1, 2$ is smaller for the variety of coal generating lesser emission for every unit used (*i.e.,* with a smaller value of $\psi_{\mathbb{c}_i}\phi_{\mathbb{c}_i}$).

Secondly, a Pigouvian subsidy equal to $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{S}}$ is due on efforts towards CCS. This is because $\frac{\bar{\gamma}_Y}{\bar{\kappa}}\psi_{\mathbb{S}}$ measures the monetised value of social marginal benefits from CCS (which implies carbon externality mitigation). This subsidy is in contrast to subsidies required to meet quantity targets that were studied in the previous two sections. Credit for CCS efforts to firms have been discussed in Metcalf (2009) and Levell, O'Connell, and Smith (2016), who suggest refunding carbon excise paid by carbon-externality generating firms who also engage in offsetting CCS activities. In addition, the analysis here supports a market for CCS services rendered by firms who may not themselves be generating carbon externalities.

8 Conclusions.

The objective of the current work was to develop a stylised model based on some key assumptions to derive some principles for the design of commodity taxes, when they serve the revenue, redistribution, and externality-correction objectives of the government, and to analyse and compare some real-life commodity tax policies against these principles.

We begin this concluding section with a few cursory remarks on recent developments in the commodity tax structure of the Indian economy, which has just switched to the GST, as these can be contrasted with the results obtained in this paper. Note that, in 2014, indirect taxation contributed to as much as 67% of the total tax revenue of the Indian government. While five basic tax rates, *viz.,* 0%, 5%, 12%, 18%, and 28% are implemented under the new GST; there is a considerable debate over how the goods are to be classified into these four slabs. Public and business outcry forced the Indian government to change its initial allocation of goods to various GST slabs, and as many as 178 goods were shifted from the 28% GST slab to the 18% slab.[50] There was also a significant reclassification of goods in the other slabs to lower slabs. This clearly indicates that the tussle between equity and efficiency objectives of the government while designing commodity taxation is at play; and that the government has not yet got this balance right.

[50] These included some modern-day necessities such as detergents, washing and cleaning preparations, shampoos, and electrical wires, cables, plugs, switches, and sockets; while some other necessities such as cement for construction purposes were retained in the 28% slab.

There is also a clear conflation of the MPRR and the Pigouvian components of many commodity taxes. For example, under the new categorisation, the government stated that the 28% GST rate[51] is reserved for *luxury* and *sin* goods, where the latter are associated with detrimental external effects. In addition, some goods are also subject to compensatory cesses, which are ostensibly levied to compensate states (India has a federal system) for any losses in revenue that they may incur during the shift to the GST. However, most goods subject to this levy are associated with significant external effects, and were already taxed at very high rates of excise in the pre-GST period. They include goods such as cigarettes and other tobacco products, aerated drinks containing sugar or other sweetening matter, and coal. Taxation of alcohol, motoring fuel such as petrol and diesel, as well as electricity was traditionally a state subject and this continues to remain so, with these goods not being included under the GST, while motoring fuel is also subject to a central excise. This has meant that firms producing these goods are not allowed tax credits on the inputs purchased. This has also meant that, while in some states such as Bihar and Gujarat, there is a total ban on liquor consumption to eliminate all alcohol related external effects, in others such as Tamil Nadu the fairly inelastic demand for liquor fetches the state government considerable tax revenue.

These trends are contrary to the results of this paper, which show that the MPRR and Pigouvian components of consumption taxes are distinct, serve different roles (the MPRR component balances equity and efficiency objectives, while the Pigouvian component regulates externality generation), and are additive at a second-best optimum. If the former is implemented as a VAT and the latter as an externality excise, then producers buying these goods as inputs at retail prices would be rebated for the MPRR component, while the rebate on the externality component would depend upon whether or not the use of these goods as inputs by firms generates externalities. Further, the paper argues that optimal VAT on goods such as electricity and motor fuel could be chosen to be zero, as the demands for these goods is indirectly derived from demands for other "VATable" goods and services such as electrical appliances and motor vehicles whose consumption or usage as inputs requires consumption or usage as inputs of goods such as electricity or motor fuels. Hence, producers purchasing electricity or motor fuel as inputs would receive no tax rebates on such purchases. On the other hand, second-best optimality requires that producers of electricity or motoring fuel should not be denied credit on the MPRR (VAT) components of the taxes paid during their purchases of inputs at retail prices. The additive feature of the VAT and excise tax components of a commodity tax at a second-best optimum is in contrast to the multiplicative structure that is often implemented in real life, where for many goods such as motoring fuel, alcohol, cigarettes, *etc.*, VAT is applied on price inclusive of the excise tax. Since producers are to receive credit on VAT paid on purchases of inputs, this structure could end up rebating producers also on the externality tax paid for use of externality-causing inputs such as motoring fuel.

The intuitive or heuristic arguments made in the policy-oriented literature on the equivalence of upstream and downstream taxation of externalities are rigorously proved in this paper under

[51] Which is one of the highest in the world.

the assumption that, along manufacturing chains that include externality-causing goods, some inputs can be non-substitutable in the production of some outputs. If this assumption were not true, then taxing an output in the manufacturing chain whose production employs an externality-causing input may not imply a reduction in the use of this input. For example, if coal was *not* non-substitutable (in terms of the definition of non-substitutability provided in this paper) in the production of cement or thermal electricity, then taxing thermal electricity or cement with an aim to control CO_2 emissions may not always lead profit maximising firms to reduce their usage of coal.[52] Thus, optimal intermediate input tax rates for many goods are not independently determined. For example, we have shown that the climate change levies in UK on thermal electricity and coal imply that such levies are also required on other coal-using industries such as cement, iron and steel, and paper and pulp. At a second-best optimum, the sum of the intermediate input tax revenues collected from various sectors located upstream of a given good, say X, when one unit of good X is produced in a manufacturing chain containing externality-causing goods, should be equal to the monetised value of the social marginal (externality) cost of producing good X.

Finally, an observation of the Indian tax system indicates the need for further research that studies second-best taxation in the presence of externalities in a fiscal-federal framework.[53]

APPENDIX

Proof of Lemma 5. The lemma is a consequence of a basic result in linear algebra: If A is a $n \times n$ matrix and $\bar{v} \in \mathbf{R}^n$ solves the linear equation system $Av = b$, then the complete set of solutions of the above linear system of equations is given by $\{\bar{v}\} + Null\,(A)$.[54]

In our case, since $\bar{\rho}$ is in the null-space of $\nabla_\rho^\top \mathfrak{y}$ whose dimension is one, $\bar{\rho}$ can be taken to be its basis: $Null\left(\nabla_\rho^\top \mathfrak{y}\right) = \{\zeta \in \mathbf{R}^{\mathfrak{N}_\rho} \mid \zeta = \kappa\bar{\rho}\ \forall\ \kappa \in \mathbf{R}\}$. Hence, given that $\bar{\sigma}$ is a solution to $(*)$ in the lemma, the set of all solutions to $(*)$ is $\left\{\zeta \in \mathbf{R}^{\mathfrak{N}_\rho} \mid \zeta = \bar{\sigma} + \kappa\bar{\rho}\ \forall\ \kappa \in \mathbf{R}\right\}$. The conclusion of the lemma follows since (19) implies that, at the second-best $\bar{\Upsilon}$, ϑ is a solution to $(*)$. ■

Proof of Theorem 6

Part (i) Let $\bar{p} \in \mathcal{P}(\bar{\rho})$ and $\bar{\eta}$ be the vector of intermediate input taxes associated with it. It follows from (23) that

$$\bar{p}_{Oi} = \bar{p}_{Ii} \implies \bar{\eta}_i = 0 \quad \forall\ i \in \widehat{\mathcal{N}}, \qquad \bar{p}_{O\mathbb{t}} = \bar{p}_{I\mathbb{ta}} \implies \bar{\eta}_{\mathbb{ta}} = 0, \qquad \bar{p}_{O\mathbb{r}} = \bar{p}_{I\mathbb{e}} \implies \bar{\eta}_{\mathbb{r}} = 0.$$

Since $\bar{\nu}_{\mathbb{c}} - \phi_{\mathbb{v}}\bar{\nu}_{\mathbb{v}} = \bar{\kappa}\bar{\rho}_{O\mathbb{c}}$, we have $\bar{\nu}_{\mathbb{c}} = \phi_{\mathbb{v}}\bar{\nu}_{\mathbb{v}} + \bar{\kappa}\bar{\rho}_{O\mathbb{c}}$. Substituting in $\bar{\nu}_{\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\bar{\nu}_{\mathbb{c}} = \bar{\kappa}\bar{\rho}_{O\mathbb{w}_i} - \bar{\gamma}_Y\psi_{\mathbb{c}}\phi_{\mathbb{cw}_i}$

[52] Rather tax-induced reductions in output production could be effected by reductions in inputs other than coal.

[53] While there is a literature that studies tax harmonisation in a fiscal federal framework, it abstracts from the problem of externalities. See *e.g.,* Blackorby and Brett (2000).

[54] See *e.g.,* Simon and Blume (1994) for proof.

for all $i = 1, \ldots, w$, we obtain

$$\begin{aligned}
\bar\kappa\bar p_{I\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\left[\phi_{\mathbb{v}}\bar\nu_{\mathbb{v}} + \bar\kappa\bar\rho_{O\mathbb{c}}\right] &= \bar\kappa\bar\rho_{O\mathbb{w}_i} + \bar\gamma_Y\psi_{\mathbb{c}}\phi_{\mathbb{cw}_i} \\
\Longrightarrow\ \bar\kappa\bar p_{I\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\left[\phi_{\mathbb{v}}\bar p_{I\mathbb{v}} + \bar\kappa\left(\bar p_{O\mathbb{c}} - \phi_{\mathbb{v}}\bar p_{I\mathbb{v}}\right)\right] &= \bar\kappa\left(\bar p_{O\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\bar p_{I\mathbb{cw}_i}\right) + \bar\gamma_Y\psi_{\mathbb{c}}\phi_{\mathbb{cw}_i} \\
\Longrightarrow\ \left[\bar p_{I\mathbb{w}_i} - \bar p_{O\mathbb{w}_i}\right] + \phi_{\mathbb{cw}_i}\left[\bar p_{I\mathbb{cw}_i} - \bar p_{O\mathbb{c}}\right] + \phi_{\mathbb{v}}\phi_{\mathbb{cw}_i}\left[\bar p_{I\mathbb{v}} - \bar p_{O\mathbb{v}}\right] &= \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cw}_i} \\
\Longrightarrow\ \bar\eta_{\mathbb{w}_i} + \phi_{\mathbb{cw}_i}\bar\eta_{\mathbb{cw}_i} + \phi_{\mathbb{v}}\phi_{\mathbb{cw}_i}\bar\eta_{\mathbb{v}} &= \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cw}_i}
\end{aligned}$$

and substituting in $\bar\nu_{\mathbb{e}} - \phi_{\mathbb{cn}}\bar\nu_{\mathbb{c}} = \bar\kappa\bar\rho_{O\mathbb{n}} + \bar\gamma_Y\psi_{\mathbb{c}}\phi_{\mathbb{cn}}$, we obtain

$$\begin{aligned}
&\bar\nu_{\mathbb{e}} - \phi_{\mathbb{cn}}\left[\phi_{\mathbb{v}}\bar\nu_{\mathbb{v}} + \bar\kappa\bar\rho_{O\mathbb{c}}\right] = \bar\kappa\bar\rho_{O\mathbb{n}} + \bar\gamma_Y\psi_{\mathbb{c}}\phi_{\mathbb{cn}} \\
&\Longrightarrow\ \bar p_{I\mathbb{e}} - \phi_{\mathbb{cn}}\left[\phi_{\mathbb{v}}\bar p_{O\mathbb{v}} + \bar p_{O\mathbb{c}} - \phi_{\mathbb{v}}\bar p_{I\mathbb{v}}\right] = \bar p_{O\mathbb{n}} - \phi_{\mathbb{cn}}\bar p_{I\mathbb{cn}} + \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cn}} \\
&\Longrightarrow\ \left[\bar p_{I\mathbb{e}} - \bar p_{O\mathbb{n}}\right] + \phi_{\mathbb{cn}}\left[\bar p_{I\mathbb{cn}} - \bar p_{O\mathbb{c}}\right] + \phi_{\mathbb{v}}\phi_{\mathbb{cn}}\left[\bar p_{I\mathbb{v}} - \bar p_{O\mathbb{v}}\right] = \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cn}} \\
&\Longrightarrow\ \bar\eta_{\mathbb{n}} + \phi_{\mathbb{cn}}\bar\eta_{\mathbb{cn}} + \phi_{\mathbb{v}}\phi_{\mathbb{cn}}\bar\eta_{\mathbb{v}} = \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cn}} \qquad (56)
\end{aligned}$$

$$\begin{aligned}
\bar\nu_{\mathbb{l}} - \phi_{\mathbb{t}}\bar\nu_{\mathbb{t}} &= \bar\kappa\bar\rho_{O\mathbb{l}} = \bar\kappa\left(\bar p_{O\mathbb{l}} - \phi_{\mathbb{t}}\bar p_{I\mathbb{tl}}\right) &&\Longrightarrow\ \bar\nu_{\mathbb{l}} - \phi_{\mathbb{t}}\bar\kappa\bar p_{O\mathbb{t}} = \bar\kappa\left(\bar p_{O\mathbb{l}} - \phi_{\mathbb{t}}\bar p_{I\mathbb{tl}}\right) \\
\Longrightarrow\ \bar\nu_{\mathbb{l}} &= \bar\kappa\bar p_{O\mathbb{l}} - \phi_{\mathbb{t}}\bar\kappa\left(\bar p_{I\mathbb{tl}} - \bar p_{O\mathbb{t}}\right) &&\Longrightarrow\ \bar\nu_{\mathbb{l}} = \bar\kappa\left(\bar p_{O\mathbb{l}} - \phi_{\mathbb{t}}\bar\eta_{\mathbb{tl}}\right) \qquad (57) \\
\Longrightarrow\ \bar\kappa\bar p_{I\mathbb{l}} &= \bar\kappa\left(\bar p_{O\mathbb{l}} - \phi_{\mathbb{t}}\bar\eta_{\mathbb{tl}}\right) &&\Longrightarrow\ \bar\eta_{\mathbb{l}} = -\phi_{\mathbb{t}}\bar\eta_{\mathbb{tl}} \qquad (58)
\end{aligned}$$

Since $\bar\nu_{\mathbb{o}} - \phi_{\mathbb{u}}\bar\nu_{\mathbb{u}} = \bar\kappa\bar\rho_{O\mathbb{o}}$, we have $\bar\nu_{\mathbb{o}} = \phi_{\mathbb{u}}\bar\nu_{\mathbb{u}} + \bar\kappa\bar\rho_{O\mathbb{o}}$. For $i = \mathbb{s}, \mathbb{m}, \mathbb{b}$, substituting in $\bar\nu_i + \alpha_i\bar\nu_{\mathbb{o}} + \beta_i\bar\nu_{\mathbb{R}} = \bar\kappa\bar\rho_{Ii} - \bar\gamma_Y\psi_{\mathbb{o}}\alpha_i - \bar\gamma_C\psi_{\mathbb{R}}\beta_i$ we obtain

$$\begin{aligned}
\bar\nu_i + \alpha_i\bar\nu_{\mathbb{o}} + \beta_i\bar\nu_{\mathbb{R}} = \bar\nu_i + \alpha_i\left[\phi_{\mathbb{u}}\bar\nu_{\mathbb{u}} + \bar\kappa\bar\rho_{O\mathbb{o}}\right] + \beta_i\bar\nu_{\mathbb{R}} &= \bar\kappa\bar\rho_{Ii} - \bar\gamma_Y\psi_{\mathbb{o}}\alpha_i - \bar\gamma_C\psi_{\mathbb{R}}\beta_i \\
\Longrightarrow\ \bar\kappa\bar p_{Oi} + \alpha_i\bar\kappa\left[\phi_{\mathbb{u}}\bar p_{O\mathbb{u}} + \left(\bar p_{O\mathbb{o}} - \phi_{\mathbb{u}}\bar p_{I\mathbb{u}}\right)\right] + \bar\kappa\beta_i\bar p_{O\mathbb{R}} &= \bar\kappa\left(\bar p_{Ii} + \alpha_i\bar p_{I\mathbb{o}} + \beta_i\bar p_{I\mathbb{R}}\right) - \bar\gamma_Y\psi_{\mathbb{o}}\alpha_i - \bar\gamma_C\psi_{\mathbb{R}}\beta_i \\
\Longrightarrow\ \bar\eta_i + \alpha_i\bar\eta_{\mathbb{o}} + \beta_i\bar\eta_{\mathbb{R}} + \alpha_i\phi_{\mathbb{u}}\bar\eta_{\mathbb{u}} &= \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{o}}\alpha_i + \frac{\bar\gamma_C}{\bar\kappa}\psi_{\mathbb{R}}\beta_i
\end{aligned}$$

Similarly, we can show that, at the second-best $\bar\Upsilon$, we have: $\bar\eta_{\mathbb{k}} + \alpha_{\mathbb{k}}\bar\eta_{\mathbb{o}} + \alpha_{\mathbb{k}}\phi_{\mathbb{u}}\bar\eta_{\mathbb{u}} = \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{o}}\alpha_{\mathbb{k}}$.
Part (ii) Let $\bar\eta \in \Omega_\eta$. At the second-best $\bar\Upsilon$, the effective output and input producer price vectors are $\bar\rho_O \in \mathbf{R}_+^{\mathfrak{N}_{\rho_O}}$ and $\bar\rho_I \in \mathbf{R}_+^{\mathfrak{N}_{\rho_I}}$. Define a vector $\bar p = \langle\bar p_O, \bar p_I\rangle \in \mathbf{R}^{\mathfrak{N}}$ with elements

$$\begin{aligned}
\bar p_{Oi} &= \bar\rho_{Oi} \quad \forall\, i \in \{\widetilde{\mathbb{N}}, \mathbb{t}, \mathbb{r}, \mathbb{v}, \mathbb{u}\}, \qquad \bar p_{Ii} = \bar\rho_{Ii} \quad \forall\, i \in \{\widehat{\mathbb{N}}, \mathbb{w}_1, \ldots, \mathbb{w}_w, \mathbb{ta}, \mathbb{e}, \mathbb{l}\}, \\
\bar p_{O\mathbb{n}} &= \phi_{\mathbb{c}}\bar\eta_{\mathbb{cn}} + \phi_{\mathbb{v}}\phi_{\mathbb{c}}\bar\eta_{\mathbb{v}} + \bar\rho_{I\mathbb{e}} - \frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}}\phi_{\mathbb{cn}}, \qquad \bar p_{I\mathbb{v}} = \bar\eta_{\mathbb{v}} + \bar\rho_{O\mathbb{v}}, \\
\bar p_{O\mathbb{c}} &= \phi_{\mathbb{v}}\bar\eta_{\mathbb{v}} + \phi_{\mathbb{v}}\bar\rho_{O\mathbb{v}} + \bar\rho_{O\mathbb{c}}, \qquad \bar p_{I\mathbb{cn}} = \bar\eta_{\mathbb{cn}} + \phi_{\mathbb{cn}}\bar\eta_{\mathbb{v}} + \phi_{\mathbb{v}}\bar\rho_{O\mathbb{v}} + \bar\rho_{O\mathbb{c}}, \\
\bar p_{I\mathbb{cw}_i} &= \bar\eta_{\mathbb{cw}_i} + \phi_{\mathbb{cw}_i}\bar\eta_{\mathbb{v}} + \phi_{\mathbb{v}}\bar\rho_{O\mathbb{v}} + \bar\rho_{O\mathbb{c}} \quad \forall\, i = 1, \ldots, w \qquad \bar p_{O\mathbb{w}_i} = \bar\rho_{I\mathbb{w}_i} - \phi_{\mathbb{cw}_i}\left(\frac{\bar\gamma_Y}{\bar\kappa}\psi_{\mathbb{c}} - \bar\eta_{\mathbb{cw}_i} - \phi_{\mathbb{v}}\bar\eta_{\mathbb{v}}\right), \\
\bar p_{O\mathbb{l}} &= \phi_{\mathbb{t}}\bar\eta_{\mathbb{tl}} + \bar\rho_{I\mathbb{l}}, \qquad \bar p_{Ii} = \bar\eta_i + \bar\rho_{Oi}, \quad \forall\, i = \mathbb{s}, \mathbb{m}, \mathbb{b}, \mathbb{k}, \qquad \bar p_{I\mathbb{u}} = \bar\eta_{\mathbb{u}} + \bar\rho_{O\mathbb{u}}, \\
\bar p_{I\mathbb{R}} &= \bar\eta_{\mathbb{R}} + \bar\rho_{O\mathbb{R}}, \qquad \bar p_{O\mathbb{o}} = \bar\rho_{O\mathbb{o}} + \phi_{\mathbb{u}}\left(\bar\eta_{\mathbb{u}} + \bar\rho_{O\mathbb{u}}\right), \qquad \bar p_{I\mathbb{o}} = \bar\eta_{\mathbb{o}} + \bar\rho_{O\mathbb{o}} + \phi_{\mathbb{u}}\left(\bar\eta_{\mathbb{u}} + \bar\rho_{O\mathbb{u}}\right)
\end{aligned}$$

If $\bar p \in \mathbf{R}_+^{\mathfrak{N}}$, then $\bar\eta$ is the associated vector of intermediate input taxes and $\bar p \in \mathcal{P}(\bar\rho)$. ■

REFERENCES

Adam, S., T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds) *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010.

Ahlberg, J. A., "Optimal Taxation of Intermediate Goods in the Presence of Externalities: A Survey Towards the Transport Sector" A Draft, 2006

Atkinson, A. B. and J. E. Stiglitz, "The Design of Tax Structure: Direct versus Indirect Taxation" *Journal of Public Economics* **6**, 1976, 55-75

Atkinson, A. B. and J. E. Stiglitz, *Lectures on Public Economics*, McGraw-Hill, 1980

Atkinson, A. B. and N. H. Stern, "Pigou, Taxation and Public Goods" *Review of Economics Studies* **41**, 1974, 119-128

Barthold T. A., "Issues in the Design of Environmental Excise Taxes" *Journal of Economic Perspectives*, **8**, 1994, 133-151

Baumgärtner, S. and J de S Arons, "Necessity and Inefficiency in the Generation of Waste: A thermodynamic Analysis" *Journal of Industrial Ecology*, **7**, 2003, 113-123

Baumol W. J. and W.E. Oates, *The Theory of Environmental Policy*, Second ed., Cambridge University Press Boyd, Cambridge, 1988.

Blackorby, C., and C. Brett, 2000. "Fiscal Federalism Revisited" *Journal of Economic Theory*, **92**, 2000, 300-317

Bovenberg, A. L. and F. van der Ploeg, "Environmental Policy, Public Finance and the Labour Market in a Second Best World" *Journal of Public Economics* **55**, 1994, 349-90

Bovenberg, A. L. and R. A. de Mooij, "Environmental Levies and Distortionary Taxation" *American Economics Review* **84**, 1994, 1085-1089

Bovenberg, A. L. and L. H. Goulder, "Optimal Environmental Taxation in Presence of Other Taxes: General Equilibrium Analyses" *American Economic Review* **84**, 1996, 985-1000

Crawford, I., M. Keen, and S. Smith, "Value Added Tax and Excises," Chapter 4 in J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds) *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010.

Dasgupta, P. and J. Stiglitz, "On Optimal Taxation and Public Production" *The Review of Economic Studies* **39**, 1972, 87-103.

De Borger, B., "Optimal Pricing of Public Final and Intermediate Goods in the Presence of Externalities," SESO Report 96/336, Universitaire-Faculteiten St. Ignatius, Antwerp

Diamond, P. and J. Mirrlees(a), "Optimal Taxation and Public Production I: Production Efficiency" *The American Economic Review* **61**, 1971, 8-27.

Diamond, P. and J. Mirrlees(b), "Optimal Taxation and Public Production II: Tax Rules" *The American Economic Review* **61**, 1971, 261-278.

Fullerton, D., A. Leicester, and S. Smith, "Environmental Taxes," Chapter 5 in S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds) *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010.

Goulder, L. H., “Environmental Taxation and the Double Dividend: A Reader’s Guide” *International Tax and Public Finance* **2**, 1995, 157-83

Guesnerie, R., *A Contribution to the Pure Theory of Taxation*, Cambridge, UK: Cambridge University Press, 1995.

Levell, P., M. O’Connell, K. Smith, “Excise Duties” Chapter 9 in C. Emmerson, P. Johnson and R. Joyce (eds) *IFS Green Budget 2016*, Institute for Fiscal Studies, 2016

Mansur, E., “Upstream versus Downstream Implementation of Climate Policy” in D/ Fullerton and C. Wolfram (eds) *The Design and Implementation of U.S. Climate Policy*, 179-193, National Bureau of Economic Research, 2012

Mayeres, I. and S. Proost, “Optimal Tax and Public Investment Rules for Congestion Type of Externalities” *The Scandinavian Journal of Economics* **99**, 1997, 261–279

Mayeres, I. and S. Proost, “Optimal Tax Rules with Congestion Type of Externalities” CES working paper, Catholic University of Leuven, Leuven, 1994

Metcalf, G. E., “Environmental Taxation: What Have We learnt in This Decade?” in A. Viard (ed) *Tax Policy. Lessons From the 2000s*, 7-34, AEI Press, 2009

Metcalf, G. E., “Designing a Carbon Tax to Reduce U. S. Greenhouse Gas Emissions” *Review of Environmental Economics and Policy* **3**, 2009, 63-83

Mirrlees, J. A., “Optimal Tax Theory: A Synthesis” *Journal of Public Economics* **6**, 1976, 327-358

Mirrlees, J., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds) *Tax by Design: the Mirrlees Review*, Oxford University Press, 2010.

Murty, S., “ Production Efficiency and Constraints on Profit Taxation and Profit Distribution in Economies with Ramsey Taxation” *Social Choice and Welfare* **41**, 2013, 579-604

Murty, S., “Lessons from Optimal Taxation for the GST in India and Beyond” *Centre for Trade and Development, School of International Studies, Jawaharlal Nehru University Discussion Paper Series* No. 17-07, 2017

Myles, G. D., *Public Economics*, Cambridge, UK: Cambridge University Press, 1995.

Ramsey, F., “A Contribution to the Theory of Taxation” *The Economic Journal* **37**, 1927, 47-61.

Sandmo, A., “Optimal Taxation in the Presence of Externalities” *The Swedish Journal of Economics* **77**, 1975, 86-98

Simon, C. P., and L. E. Blume, *Mathematics for Economists* W.W. Norton, London 1994

Weymark J., “A Reconciliation of Recent Results in Optimal Taxation Theory” *Journal of Public Economics* **12**, 1979, 171-189.

Williams III, R. C., “Environmental Taxation” *NBER Working Paper 22303,* National Bureau of Economic Research, Cambridge, MA, 2016