TAXING EXTERNALITIES WITHOUT HURTING THE POOR

By

Mallesh Pai and Philipp Strack

August 2022

COWLES FOUNDATION DISCUSSION PAPER NO. 2377

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

TAXING EXTERNALITIES WITHOUT HURTING THE POOR

MALLESH M. PAI[A] AND PHILIPP STRACK[B]

ABSTRACT: We consider the optimal taxation of a good which exhibits a negative externality, in a setting where agents differ in their value for the good, their disutility from the externality, and their value for money, while the planner observes neither. Pigouvian taxation is the unique Pareto efficient mechanism, yet it is only optimal if the planner puts higher Pareto weights on richer agents. We derive the optimal tax schedule for both a narrow allocative objective and a utilitarian objective for the planner. The optimal tax is generically nonlinear, and Pareto inefficient. The optimal mechanism might take a "non-market" form and cap consumption, or forbid it altogether. We illustrate the tractability of our model by deriving closed form solutions for the lognormal and Rayleigh distribution. Finally, we calibrate our model and derive optimal taxes for the case of air travel.

KEYWORDS: externalities, redistribution, taxation, mechanism design.

JEL CLASSIFICATION: D82, H23.

[A]DEPARTMENT OF ECONOMICS, RICE UNIVERSITY, MALLESH.PAI@RICE.EDU
[B]DEPARTMENT OF ECONOMICS, YALE UNIVERSITY, PHILIPP.STRACK@YALE.EDU
JANUARY 24, 2023
We thank Mohammad Akbarpour, Itai Ashlagi, Laura Doval, Piotr Dworczak, Paul Heidhues, Ravi Jagadeesan, Zi Yang Kang, Botond Koszegi, Shengwu Li, Meg Meyer, Frikk Nesje, Mike Ostrovsky, Andrea Prat, Maher Said, Ilya Segal, Klaus Schmidt, Vasiliki Skreta, Andy Skrzypacz, Steve Tadelis, Dmitry Taubinsky, Alex Wolitzky and various seminar audiences for insightful comments and discussions. We thank Neil He, Dibya Mishra and Tracy Zhou for excellent research assistance. Pai gratefully acknolwedges support from the NSF (CCF-1763349).

1. INTRODUCTION

When there are differences between the private and social costs of a good, i.e., when consumption of the good causes externalities, market outcomes may be inefficient. Economists have long recognized this market failure and suggested corrective taxation as a remedy: Pigou (1920) proposed a tax on polluters equal to the monetary equivalent of the harm done to others.[1] Such a tax leads polluters to internalize the externality they cause and thereby reestablishes efficiency.

Despite their simplicity and intuitive appeal, Pigouvian taxes are rarely seen in practice. One reason could be that agents' values for money may be heterogeneous due to wealth and income effects. A Pigouvian tax may further disadvantage people with lower income: e.g., a gasoline tax might disproportionately exclude people of low income from driving cars and thereby severely limit their life. Furthermore, (and potentially worse) it might have little effect on people of high income and low value for money if their demand for the good reacts relatively little to the price change induced by the tax. Conversely, public discourse often involves regulations such as consumption caps or even prohibition of the activity, which economists consider "non-market" solutions. Are the former concerns well-founded, and can natural preferences of the planner justify the latter?

To answer these questions, we study a setting where a continuum of agents must each decide their consumption of a good. Agents receive a private benefit from their consumption and suffer an externality cost that depends on the total consumption in society. Agents have heterogeneous utility from consumption, externality costs and value for money. All of these are privately known to the agent, i.e., this is a setting of multi-dimensional private information. Formally, our setting is in the tractable class of linear preference models explored by Condorelli (2013), Akbarpour et al. (2020), and Dworczak et al. (2021), described in more detail below.

A planner would like to regulate consumption in this setting. Since the planner does not observe agents' private information, they can only tax consumption and redistribute the proceeds. We leverage the tractability of our model to derive novel economics regarding the taxation of externalities. We consider three natural objectives for the planner, namely: Pareto efficiency, utilitarian efficiency, and narrow allocative efficiency.

The first two objectives are standard and need no further introduction. We believe narrow allocative efficiency is novel in the public finance literature on the regulation of externalities and therefore warrants further discussion. To understand the motivation for this objective, note that in a setting where agents have different values for money, regulation of an externality by a planner can be beneficial in two ways: first, of course, this may reduce the amount of externalities suffered by agents. Secondly, excess revenues raised

[1] For a framing of Pigou's ideas in modern economic terms see Baumol (1972).

can potentially be redistributed to agents with a higher value for money. A planner with a narrow allocative efficiency objective is concerned with the net social surplus generated *only* by the allocation. Such a planner cares purely about ensuring agents consume the good if and only if they value it more than the externalities and costs its consumption creates for society, ignoring any welfare impacts of redistributive transfers. By contrast, utilitarian efficiency takes into account the gains from redistribution. To realize these gains, a utilitarian planner may be willing to tolerate some inefficiency in the allocation of the good, for example, by having rich people who value the good less than its externality consume it simply to raise money to distribute to the poor.

Whether the planner should consider redistribution when designing policies that regulate the consumption of goods with externalities (e.g. air travel) is a philosophical question and the subject of an active political debate. As a case study, consider the public debate on the "Green New Deal" in the US.[2] It was proposed by members of the Democratic party in 2019, with various regulations to control pollution/greenhouse gas emitting activities and their associated climate externalities. However, it also proposed using these revenues towards social programs (higher education, health care) which would benefit the poorer sections of society. This led to criticism for mixing redistribution elements with climate protection: for example, the Republican party sponsored a resolution in the US house of commons that in part criticized the Green New Deal for " *the enactment of policies with no semblance of a connection to emissions or climate change.*"[3]

Allocative and utilitarian efficiency, therefore, take the two extreme stances on how much a planner should account for the value of redistribution. In the following analysis, we capture both the allocative and utilitarian perspectives, as well as any convex combination of the two.

When agents have quasilinear preferences and equal values for money, all three of these objectives are equivalent. However, since agents have private and heterogeneous values for money, an allocatively efficient tax scheme that is also budget-balanced (i.e., redistributes all the tax revenues) might be neither utilitarian nor Pareto efficient.

First, we show that there exists a unique Pareto efficient outcome that is implementable despite the asymmetric information (Proposition 1).[4] This outcome is uniquely implemented by Pigouvian taxation: the planner sets a linear tax/price on consumption that equals the monetary societal cost of a marginal increase in the level of externality. As is well known, a Pareto efficient outcome can equivalently be thought of as the planner

[2]See, e.g., `https://en.wikipedia.org/wiki/Green_New_Deal`.

[3]See `https://mikejohnson.house.gov/sites/mikejohnson.house.gov/files/johnson%20green%20new%20deal%20response%20resolution.pdf`.

[4]The outcome is unique if the agent's utilities are strictly concave.

maximizing an objective function where individual agents' utilities are weighted according to some given *Pareto weights*. We show that the outcome from Pigouvian taxation corresponds to a planner who puts *higher* Pareto weights on agents with low values for money, i.e., intuitively, cares *more* about the welfare of richer agents. For example, assuming log-utility, Pigouvian taxes are only optimal for a planner who cares 10 times more about an agent who is 10 times richer, or, to put it differently, cares 27 times more about the top 10% of US society in terms of household wealth than about the bottom 50% .[5] Similarly, we establish that Pigouvian taxes are higher for externalities that mostly affect the rich (keeping the overall externality on society fixed), whereas the tax that is optimal for a planner with a utilitarian or allocative objective only depends on the overall externality. These results may formally explain the objection commonly raised in public discourse that Pigouvian taxes favor the rich. As the Pigouvian tax is the unique Pareto efficient incentive compatible mechanism, it is impossible to achieve Pareto efficiency without assigning higher Pareto weights to agents with a low value for money.

This observation extends beyond Pareto efficiency. We consider a weaker notion of *aggregate-consumption-constrained Pareto Efficiency*, i.e., Pareto efficiency subject to a fixed total level of externality causing activity, and show that this is also uniquely implemented by a linear tax (with a different rate from the Pigouvian one).

We then consider the problem where Pareto efficiency is not required and fully characterize the solution to the planner's problem for any convex combination of the narrow allocative efficiency objective and the utilitarian objective. The planner's problem can be written as a multidimensional screening problem. Incentive compatibility limits the feasible set of allocations for the planner: A planner cannot distinguish between agents who have a high value for the good and a high value for money, from ones who have a proportionally lower value for the good and lower value for money. Therefore, the transfer and allocation of an agent can only depend on the ratio of these quantities, which we term their willingness to pay (Proposition 2). Standard mechanism design techniques allow us to then rewrite the seller's problem as purely one of maximizing a virtual surplus (Proposition 3). We establish that the virtual surplus is the weighted sum of (i) an agent's value for the good conditional on willingness to pay (WTP) (ii) the average value for money of agents with a higher WTP and (iii) the standard Myersonian virtual value. This virtual surplus (7) has a natural interpretation in the context of our model.

We show that the outcomes implemented by the optimal mechanism for the objectives above are, generally, not aggregate-consumption-constrained Pareto efficient. In particular, we show that a planner with such motives will generically impose a *nonlinear* tax,

[5] In Q4 2021, the top 10% of households in the United States held 69.4% of the country's wealth, while the bottom 50% held 2.6% (The Federal Reserve, 2022).

i.e., the marginal tax on consumption is not constant (Proposition 4).[6] Further, this means that a planner concerned with any (convex combination) of these objectives cannot allow resale of the good (which implies aggregate-consumption-constrained Pareto Efficiency). We also qualitatively describe how the optimal mechanism differs from Pigouvian taxation depending on the designer's objective. A planner with a utilitarian objective will induce lower consumption of the poor and higher consumption of the rich as doing so increases the revenue raised by taxes and thus allows for higher transfers to the poor (Proposition 5). In sharp contrast, a planner who is interested in narrow allocative efficiency will increase the consumption of the poor as their willingness to pay understates their benefit of consumption and decrease the consumption of the rich as their willingness to pay overstates their benefit of consumption (Proposition 6).

To demonstrate the tractability of our formulation and results, we consider some special cases for the multidimensional distribution of types in society. We obtain explicit analytic functional forms for the virtual value when both value for money and value for the good are independently distributed according to the Rayleigh distribution (Proposition 7) or jointly-lognormally distributed (Proposition 8). We use this to show explicit examples of settings where consumption caps, complete shutdown etc can be the optimal mechanism for the planner. Further, we show additional qualitative results, in particular the shape of the marginal tax that implements the optimal mechanism. Most intriguingly, we show that in the lognormal case, the optimal for a utilitarian planner *always* involves a *decreasing* marginal tax, i.e. the tax is decreasing in consumption (Proposition 8). In particular, the high initial tax excludes low WTP types, and society as a whole is better served by redistribution of the tax proceeds and lower overall externality-causing activity. We have not seen decreasing marginal taxes discussed in the literature; nevertheless, it arises here for a natural objective function for the planner (utilitarian) and a natural distribution (lognormal distribution), which is, in our opinion, something of a puzzle. Further, our results show how even basic properties of the optimal mechanism depend on the shape of the distribution, not just summary statistics: in the Rayleigh distribution setting, a utilitarian planner would charge an *increasing* marginal tax (Proposition 7).

As an illustration, Section 8 applies our results to the context of air travel. Using available air-travel data, and plausible preferences of the planner, we calibrate the lognormal model described above and calculate the optimal mechanism for the allocative and the utilitarian objectives, and compare these to Pigouvian taxation. In a setting where the Pigouvian planner would charge a flat tax of \$50 per flight, we show that utilitarian optimal taxes are decreasing and $\approx$ \$200 per flight, while allocative optimal taxes are increasing and range from $\approx$ \$125 for the first flight to over \$200 for the sixth flight an individual

[6]We discuss the possibility of implementing such a tax in practice in Section 9.

takes. Both are therefore substantially higher than the Pigouvian tax and lead to higher rebates to households (≈ \$750–\$1035 versus \$280 in the Pigouvian case). As the majority of the US population does not fly in a given year such taxes would plausibly make the majority of the population significantly better off.

1.1. *Related Literature*

Before we proceed, a brief review of the related literature is useful. The tractability of our modeling owes in part to the idea that agents' have linear preferences over transfers, but different values for money, or equivalently, that the planner's preferences are such that he values redistribution of money between the agents.[7]

In mechanism design, this was pioneered by Condorelli (2013), who in the context of allocation problems gave conditions under which the optimal mechanism for a planner may have "non-market mechanism" features. A key recent paper is the work of Akbarpour et al. (2020), who study trade between buyers and sellers. They show that the optimal mechanism depends on the nature of inequality in society, i.e., within group (e.g., buyers are heterogeneous in their value for money) or across group (e.g., buyers value money differently from sellers). Various mechanisms disfavored by "traditional" economic arguments, like price ceilings or floors, can be optimal. Dworczak et al. (2021) consider the allocation of a good of heterogeneous quality when buyers have quasilinear preferences and private values for quality, and the planner assigns different welfare weights. The optimality of market vs non-market mechanisms depends on the correlation between agents' values for quality and the planner's welfare weights. Akbarpour et al. (2022) apply this framework to the question of allocating scarce vaccines and identify conditions under which a planner prefers prices over priorities assigned based on observable characteristics (and vice versa). Closest to our analysis is the contemporaneous work of Kang (2022) which also considers the optimal taxation of individuals who produce heterogeneous externalities via an imperfect proxy (consumption). In contrast, we restrict attention to homogeneous externalities, but agents care differently about the externality. Assuming an allocative objective slightly different from ours, Kang identifies conditions such that quantity floors and ceilings are optimal and quantifies the benefits of nonlinear taxation.

In the public finance literature, after the seminal work of Pigou (1920) there has been a long tradition of studying corrective taxation (see e.g., Baumol, 1972; Diamond, 1973). Tullock (1967) suggested that there is an "excess benefit" of such taxes as the revenues raised from such taxes may allow the planner to reduce other standard taxes and the

[7]These welfare weights can be interpreted as marginal welfare weights in the sense of Saez and Stantcheva (2016).

associated economic distortions.[8] This hypothesis was (inconclusively) empirically evaluated by Fullerton and Metcalf (1997).

We contribute to the study of Pigouvian taxes by characterizing under what conditions Pigouvian taxes are optimal for a social planner. Maybe surprisingly, we find that Pigouvian taxes are regressive in the sense that they are only optimal for a utilitarian planner who assigns higher welfare weights to the rich and has no preference for redistribution. In addition to the utilitarian benchmark considered by most of the public finance literature, we propose "allocative efficiency" as an objective that purely cares about the allocation of the good without taking into account transfers.

Concerns about the regressiveness of externality and internality correcting taxes have also been raised in the public finance literature—for example, on cigarette taxes (Goldin and Homonoff, 2013), sugary soda taxes (Allcott et al., 2019) and taxes on energy consumption/ energy efficiency subsidies (Allcott et al., 2015; Borenstein and Davis, 2016; Davis and Knittel, 2019). This concern is amplified for goods that are mostly consumed by the poor. For a recent treatment of nonlinear corrective taxation in the presence of redistributive motives see Section 5.2 in Ferey et al. (2022). This literature traditionally characterizes the optimal tax based on variational arguments, and it connects features of the derived scheme to sufficient statistics that may be estimated from the data (e.g. various elasticities). In the context of carbon taxes, Herweg and Schmidt (2022) compare quantity and price regulation under political constraints for consumers with moral concerns.

Less related, there is a literature on the efficiency of commodity taxation in a richer model with labor choice and income taxation. The seminal results of Atkinson and Stiglitz (1976) provide a setting in which it is inefficient to tax commodities and any revenue needs of the government are best raised by income taxation. However, as was appreciated by contemporaneous work (Mirrlees, 1976) and emphasized by the authors themselves (Stiglitz, 2018), the assumptions underpinning the Atkinson-Stiglitz Theorem are strong, limiting its relevance for the design of tax policy. A series of influential papers then provide conditions under which commodity taxation may indeed be useful, see, e.g., Cremer et al. (2001) or Saez (2002). Note that these papers consider the taxation of normal consumption goods, so there are no possible gains from regulating an externality. When an externality causing activity is present that can be directly taxed, the so called "Principle of Targeting" (Dixit, 1985) suggests that under general conditions, optimal taxation involves Pigouvian taxation of the externality causing good, coupled with optimal taxation of the rest ignoring the externality but taking into account revenue raised by the externality tax (originally due to Sandmo (1975), the most general version we are aware of is due to Kopczuk (2003)). It is unclear if this separability applies to our model when

[8] Alternately, it may be possible to redistribute the revenues raised so as to be Pareto improving, see Sallee (2019).

also considering e.g., income taxation (see Rothschild and Scheuer (2016) for an example of a paper which combines optimal redistributive and corrective/ Pigouvian taxes): we leave this to future work.

2. The Model

There is a good which imposes an externality. The cost of producing the good is linear in the aggregate consumption level, with a cost $\kappa \geq 0$ per unit of the good consumed in society. Individuals may choose a level of consumption in some set $X \subseteq \mathbb{R}_+$ which represents the range of feasible allocations—depending on the application, it may be bounded or unbounded.

2.1. *Preferences*

There is a continuum of agents with three-dimensional types (θ, c, η) distributed according to the distribution F, with continuous joint density f, and marginals F_θ, F_c, F_η with $\theta \in [0, \overline{\theta})$, $c \in [0, \overline{c})$, and $\eta \in [\underline{\eta}, \overline{\eta}) \subseteq \mathbb{R}_{++}$. We denote by $T = [0, \overline{\theta}) \times [0, \overline{c}) \times [\underline{\eta}, \overline{\eta})$ the possible types of the agents.[9]

Consider an agent of type (θ, c, η). The value of θ determines the utility from consuming the good, c determines the disutility the agent suffers from the externality, and η determines the agent's utility of money relative to consumption of the good and the externality. Given a level of consumption $x \in X$, a transfer t paid by the agent, and an aggregate consumption of $\bar{x}$, the agent's utility equals[10]

$$\theta\, v(x) - c\, d(\bar{x}) - \eta\, t\,. \tag{1}$$

Consumption Utility: The first term captures the agent's utility from consumption of the good. Here $v : X \to \mathbb{R}_+$ is a differentiable, strictly increasing, weakly concave function. Without loss we normalize the model so that $v'(0) = 1$.[11] Thus, θ parametrizes the agent's marginal value for the first unit of the good.

Externality Cost: The second term captures the disutility suffered by the agent from aggregate consumption in society. The *damage* $d : \mathbb{R}_+ \to \mathbb{R}_+$ is a differentiable, strictly increasing, weakly convex function capturing a (possible) nonlinear relationship between the aggregate level of consumption in society and the disutility. Again, we normalize

[9] Our major focus is on activities that cause negative externalities, so we consider $c \in \Re_+$. However, as will be clear, our techniques apply regardless of whether the activity causes a negative or a positive externality. In the sequel, we also consider activities that have a positive externality.

[10] Allen and Rehbeck (2021) study when (θ, η) can/cannot be jointly identified. If (θ, η) can be jointly identified, we can additionally identify c by eliciting the agent's willingness to pay for a reduction in aggregate consumption/ resulting externality.

[11] To achieve this normalization we can redefine v as $v(\cdot)/v'(0)$ and θ as $\theta \times v'(0)$.

$d'(0) = 1$ such that c parametrizes the agent's marginal cost/ disutility from the first unit of aggregate consumption.

Value for Money: The value η parametrizes the agent's utility for money relative to her consumption utility and disutility from the externality. Without loss we normalize the average value for money in the population to 1, i.e. $\int \eta dF_\eta(\eta) = 1$. This normalization implies that the designer's utility increases/decreases by 1 unit if he increases/decreases each agent's income by 1 unit.

2.2. *Discussion*

Note that any positive affine transformation of the above utility function represents the same preference of the agent. The utility representation chosen in (1) is determined by how much the designer values redistribution of money between different agents as this determines the value of η. As is well known from Saez and Stantcheva (2016), and formally derived in our setting in Section 7, the heterogeneity in the designers preference for allocating money to different agents can be derived as a reduced-form for heterogeneity in wealth/income. In this case, agents with a high value for money will be poor and those with a low value for money will be rich—typically the designer will want to transfer money from the rich to the poor. For expositional clarity, we thus often refer to agents with a high value for money as *poor*, and agents with a low value for money as *rich*. A mathematically equivalent, but economically more general interpretation in the framework of Saez and Stantcheva (2016) is that η represents the "Generalized Social Marginal Welfare Weights". These weights can capture additional considerations of the planner (other than income) affecting the deservingness of an agent.

In a utilitarian framework, the values of η describe the designers preference for redistribution and can not be inferred from the agents' preferences and, as is common in the literature, we assume them to be given. Our results thus characterize the optimal tax scheme for a given objective and given redistributive preference of the planner. Similarly, like the literature on redistributive taxation, we assume that the η's are private information of the individual. All other parameters of the model can be identified from the agents preferences as we also illustrate in a concrete application in Section 8.

2.3. *Mechanisms*

A (direct revelation) *mechanism* specifies an allocation $x : T \to X$ of the good to agents of different types and a transfer $t : T \to \mathbb{R}$ made by the agents.[12] Given a mechanism

[12] As there is a continuum of agents it is without loss to assume that (x, t) depend only on the agent's own type.

(x,t) we denote the aggregate level of consumption in society by $\bar{x}$, i.e.,[13]

$$\bar{x} = \int x(\theta,c,\eta)dF.$$

Given a mechanism (x,t) the utility of an agent of type (θ,c,η) equals

$$U(\theta,c,\eta;x,t) = \theta\, v(x(\theta,c,\eta)) - c\, d\,(\bar{x}) - \eta\, t(\theta,c,\eta)\,. \tag{U}$$

A mechanism (x,t) is *budget feasible* if it does not run a deficit, which means that the revenue it raises exceeds the production cost $\kappa\bar{x}$ of the good

$$\int t(\theta,c,\eta)dF \geq \kappa \int x(\theta,c,\eta)dF\,. \tag{B}$$

Budget feasibility restricts mechanisms to have no outside funding, it is an immediate extension of the model to consider settings where the mechanism can either run some deficits or is required to run a given surplus (as this simply amounts to a change of κ).

A mechanism is *incentive compatible* if no type (θ,c,η) gains from misrepresenting their type as any other type (θ',c',η'), i.e.,

$$\theta\, v(x(\theta,c,\eta)) - \eta\, t(\theta,c,\eta) \geq \theta\, v(x(\theta',c',\eta')) - \eta\, t(\theta',c',\eta')\,. \tag{IC}$$

The private externality cost plays no role for incentive compatibility—due to our continuum assumption no agent affects aggregate consumption of the good.

2.4. *Welfare*

We consider three different welfare criteria:

Pareto Efficiency: A mechanism (or outcome) (x,t) is Pareto efficient if there does not exist another budget feasible mechanism (x',t') in which *every* agent is weakly better off and at least a positive measure of agents is strictly better off, i.e., there does not exist a budget feasible mechanism (x',t') s.t. for all types (θ,c,η)

$$U(\theta,c,\eta;x,t) \leq U(\theta,c,\eta;x',t')\,,$$

with the inequality being strict on a set to which F assigns positive measure.[14]

We also consider a notion of *aggregate-consumption-constrained Pareto Efficiency*. A mechanism (x,t) is aggregate-consumption-constrained Pareto efficient if there does not exist

[13]For simplicity and as there is no risk of confusion we throughout denote by $\int \cdot dF = \int_{(\theta,c,\eta)\in T} \cdot dF(\theta,c,\eta)$ the integral of F over its domain T.

[14]Note that our notion of Pareto efficiency is demanding as it requires Pareto efficiency with respect to outcomes which can not be implemented in an incentive compatible way. At the same time it only requires a Pareto undominated outcome for the actual distribution of types F instead of for all distributions (and potential reports of the agents).

another budget feasible mechanism (x', t') with the same level of aggregate consumption $\int x'(\theta, c, \eta)dF = \bar{x}$ that Pareto dominates it. This notion captures whether a given outcome (x, t) can be improved upon in the Pareto sense, fixing the level of aggregate consumption $\bar{x}$ in society—if not, then any further Pareto gains can only be achieved by changing the aggregate consumption level.

Narrow Allocative Objective: A planner concerned with narrow allocative efficiency is concerned with the net social surplus generated *only* by the allocation, ignoring any welfare impacts of transfers and who is making them. Formally, such a planner is concerned with

$$V_A(x, t) = \int \theta v(x(\theta, c, \eta)) - cd(\bar{x}) - \kappa x(\theta, c, \eta)dF.$$

An outcome (x, t) is *narrow allocative efficient* if it maximizes the above objective.

Utilitarian Objective: A planner concerned with utilitarian welfare simply cares about the total utility of the agents from the mechanism, i.e.

$$V_U(x, t) = \int U(\theta, c, \eta; x, t)dF.$$

An outcome (x, t) is *utilitarian efficient* if it maximizes the above objective.

For generality, we also consider a social planner who aims at maximizing a weighted sum of allocative and utilitarian objectives with some fixed weight $\alpha \in [0, 1]$ on the former and $1 - \alpha$ on the latter:

$$\alpha V_A(x, t) + (1 - \alpha)V_U(x, t)\,. \tag{2}$$

3. Benchmark Mechanisms

In this section, we describe three natural mechanisms that have received attention in practice and in the previous literature on corrective taxation, namely, Laissez-faire, prohibition, and Pigouvian taxation. These mechanisms are natural benchmarks to compare to.

Laissez-faire The first mechanism is simply not regulating the consumption of the good, which we call the "Laissez-faire" mechanism. In a competitive market, the good sells at the cost of production, ignoring the externality that results from its consumption.

Definition 1 (Laissez-faire)**.** *In the* Laissez-faire *(or no regulation) mechanism, the good is available at cost ($t(\theta, c, \eta) \equiv \kappa x(\theta, c, \eta)$). Agents consume the good whenever their willingness to pay $\rho = \theta/\eta$ exceeds the cost κ. For any agent, their consumption solves $x(\theta, c, \eta) \in \operatorname{argmax}_{z \in X} \theta v(z) - \eta \kappa z$.*

Prohibition Another natural mechanism is to prohibit the consumption of the good.

DEFINITION 2 (Prohibition)**.** *Under* prohibition*, there are no transfers and no agent consumes the good, i.e.,* $t, x \equiv 0$*.*[15]

While prohibiting consumption of the good seems extreme, doing so is trivially an optimal mechanism if it is allocatively efficient to do so, i.e. $\overline{\theta} \leq c$, and the designer cares only about the allocative objective, i.e. $\alpha = 1$, as in this case it yields the first-best outcome for the designer.

Furthermore, consumption/ production bans on certain goods have been used in practice. Examples include the ban on the sale of incandescent lightbulbs (in favor of energy-efficient alternatives),[16] bans on the use of certain ozone-depleting substances that were implemented worldwide in the 90s,[17] or bans on the use of certain single-use plastics such as plastic bags in several cities and countries.[18]

Pigouvian Taxes Another mechanism often discussed (especially among economists) is a Pigouvian tax (Pigou, 1920). In this mechanism, agents are free to consume the good, but pay a tax equal to the monetary externality caused by their consumption. In line with the suggestions in the literature on excess benefits (e.g., Tullock, 1967), we assume that the additional taxes collected are rebated uniformly to the agents.[19] To compute the monetary externality of consumption, note that an agent with value for money η would require a payment of c/η to be compensated for a marginal increase in the average consumption $\bar{x}$. The total payment necessary to compensate all agents for such a marginal increase of $\bar{x}$ thus equals

$$\mathbb{E}\left[c/\eta\right] d'(\bar{x}) .$$

This leads to the following definition of Pigouvian taxes:

DEFINITION 3 (Pigouvian tax)**.** *A* Pigouvian *tax charges agents the monetary externality* $\mathbb{E}\left[c/\eta\right] d'(\bar{x})x$ *of their consumption and the revenue is distributed equally, i.e.*

$$t(\theta, c, \eta) = \left(\kappa + \mathbb{E}\left[c/\eta\right] d'(\bar{x})\right) x(\theta, c, \eta) - \mathbb{E}\left[c/\eta\right] d'(\bar{x})\bar{x} .$$

[15] Note, that this mechanism can also be implemented by charging a price that exceeds the maximal willingness to pay of the agents.

[16] https://en.wikipedia.org/wiki/Phase-out_of_incandescent_light_bulbs

[17] https://www.epa.gov/ods-phaseout

[18] https://en.wikipedia.org/wiki/Plastic_bag_ban

[19] Formally, that literature pointed out that the revenues from such taxation could be used by the government, reducing the amount that needed to be raised, and therefore reducing distortions that resulted from, e.g., income taxation. Here the planner has no overall revenue needs, so any revenues raised are rebated directly to the population. As we show in Observation 2 such uniform redistribution is optimal for a planner who puts non-zero weight on the agents' utilities.

Agents consume the good whenever their willingness to pay exceeds $\kappa + \mathbb{E}[c/\eta]\, d'(\bar{x})$ *and their consumption solves*

$$x(\theta, c, \eta) \in \underset{z \in X}{\operatorname{argmax}}\, \theta v(z) - \eta(\kappa + \mathbb{E}[c/\eta]\, d'(\bar{x}))z\,. \tag{3}$$

4. A Discussion of Pigouvian Taxes and Pareto Efficiency

Given their importance, we begin by deriving various properties of Pigouvian taxes in our setting and relating them to Pareto efficiency. Suppose that c and η are independently distributed in the population, such that the externality of the good affects rich and poor equally. As $1/\eta$ is a convex function of η and $\mathbb{E}[\eta] = 1$, a Pigouvian tax will exceed the disutility c caused by the externality $\mathbb{E}[c/\eta] > \mathbb{E}[c]$. In the non-independent case, as

$$\mathbb{E}[c/\eta] = \mathbb{E}[c]\,\mathbb{E}[1/\eta] + \mathrm{Cov}(c, 1/\eta)$$

the Pigouvian tax is higher for goods whose externality affects the rich more strongly (i.e. when $\mathrm{Cov}(c, 1/\eta)$ is greater, but $\mathbb{E}[c]$ and $\mathbb{E}[1/\eta]$ are the same).

Similarly, if η and c are independent, the Pigouvian tax increases if the values for money of different agents become more diverse (in the sense of a mean-preserving spread), even though the average value for money remains fixed at 1. Intuitively, it requires a larger transfer to compensate rich agents with a low value for money for an increase in average consumption. Thus, in more unequal societies or for goods whose consumption mostly hurts the rich (keeping all else equal) Pigouvian taxes are higher.

When all agents assign the same value to money, i.e. $\eta = 1$ a.s., the Pigouvian tax is closely related to the celebrated Vickrey-Clark-Groves mechanism and is well known to be Pareto efficient, utilitarian efficient, and allocatively efficient.[20]

Observation 1. *If* $\eta = 1$ *a.s. then the Pigouvian tax (i) is Pareto efficient, (ii) allocative efficient, and (iii) utilitarian efficient.*

Pigouvian taxes are still Pareto efficient if the agents have different values for money. This follows as dividing each agent's utility function by η leaves their preferences and thus the set of Pareto efficient allocations unchanged. A planner who assigns a corresponding Pareto weight equal to the inverse of an agent's value for money would therefore maximize

$$(x, t) \in \underset{(\tilde{x}, \tilde{t})}{\operatorname{argmax}} \int \frac{1}{\eta}\, U(\theta, c, \eta; \tilde{x}, \tilde{t}) dF\,. \tag{4}$$

As Pigouvian taxes are utilitarian efficient with respect to these Pareto weights we obtain the following result:

[20]We provide a proof of this observation in the Appendix.

PROPOSITION 1 (Pareto Efficiency)**.** *The following are equivalent:*

(i) (x, t) *is the Pigouvian tax with equal redistribution.*

(ii) (x, t) *is incentive compatible and Pareto efficient.*

(iii) (x, t) *is optimal for a utilitarian planner with Pareto weights* $1/\eta$*, i.e., maximizes* (4) *subject to* (B) *and* (IC)*.*

The above proposition thus identifies Pigouvian taxes as the unique Pareto efficient and incentive compatible mechanism. It further shows that Pigouvian taxes maximize the sum of the agents' utility weighted by $1/\eta$.[21]. Pigouvian taxes and thus any Pareto efficient mechanism put *higher* weight on the utility of agents with lower values for money (e.g., rich agents). This is in line with the previous discussion that Pigouvian taxes are higher for goods whose consumption mostly hurts the rich. It formalizes the intuition commonly expressed in public writing that Pigouvian taxes favor the rich. This might be an important concern, as for example the carbon emission tax on gasoline was the main cause of protest of the Yellow Vest protests in France in 2018 (see e.g. Dorfman, 2018):

> *A carbon tax is designed to make products that emit carbon more expensive by an amount proportional to the emissions released to the atmosphere. [..] It is, by definition, fair because everyone is paying based on their individual carbon footprints. [..] The general belief, however, is that carbon taxes will be regressive and that such a situation is neither tolerable or politically sustainable.*

We note that the Pareto weights $1/\eta$ exactly offset the planner's preference for redistribution among the agents, and hence part *(iii)* of Proposition 1 implies that a Pareto efficient mechanism is optimal for a utilitarian planner if and only if the planner has no preference for redistribution. The restriction to Pareto efficient outcomes may not, therefore, be subjectively appealing.

Proposition 1 can be straightforwardly generalized to aggregate-consumption-constrained Pareto Efficiency.

COROLLARY 1 (aggregate-consumption-constrained Pareto Efficiency)**.** *The following are equivalent:*

(i) (x, t) *is a linear tax mechanism with equal redistribution.*

(ii) (x, t) *is incentive compatible and aggregate-consumption-constrained Pareto efficient.*

(iii) (x, t) *is optimal for a utilitarian planner with Pareto weight* $1/\eta$ *over all mechanisms that induce aggregate consumption* $\bar{x}$*, i.e. maximizes* (4) *subject to* (B) *and* (IC)*.*

[21] A related result is Negeshi's theorem, which in a general equilibrium framework *without externalities* establishes conditions such that an allocation transfer pair is an equilibrium if and only if it maximizes the weighted sum of the agent's utilites, with weights equal to the inverse of their marginal value for money Negishi (1960).

This corollary follows straightforwardly from Proposition 1 and generalizes its insights. In particular, *any* linear tax is consistent with the planner having a (aggregate-consumption-constrained) Pareto efficiency motive. Further, *any* linear tax is consistent with the planner having no redistribution motive, or, alternately, placing Pareto weights that exactly offset the original preference for redistribution.

5. Optimal Mechanisms

In the previous section, we showed that a Pareto efficient outcome is implementable by a linear tax. We now characterize the optimal mechanism for a planner concerned with narrow allocative efficiency or utilitarian efficiency.

Planner's problem The planner's problem for a weight $\alpha \in [0,1]$ on narrow allocative efficiency and $(1-\alpha)$ on utilitarian efficiency is given as:

$$\max_{(x,t)} \alpha V_A(x,t) + (1-\alpha) V_U(x,t)\,, \tag{OBJ}$$

$$\text{s.t.} \int t(\theta,c,\eta) dF \geq \kappa\, \bar{x}, \tag{B}$$

$$\theta\, v(x(\theta,c,\eta)) - \eta\, t(\theta,c,\eta) \geq \theta\, v(x(\theta',c',\eta')) - \eta\, t(\theta',c',\eta')\,. \tag{IC}$$

We begin by characterizing incentive compatible mechanisms. An agent with type (θ,c,η) has a utility of

$$\eta \left(\frac{\theta}{\eta} v(x(\theta,c,\eta)) - t(\theta,c,\eta) \right) - cd(\bar{x}).$$

The disutility caused by the externality $cd(\bar{x})$ depends only on the aggregate consumption level and thus cannot be influenced by a single agent. The agent's preferences over individual consumption transfer pairs (x,t) thus depend only on the ratio

$$\rho = \frac{\theta}{\eta} \tag{WTP}$$

of the agent's valuation for the good relative to his valuation for money, which we termed as the agent's *willingness to pay*.

Since agents' preferences only depend on their willingness to pay ρ, it is natural to conjecture that the allocation and transfer in any incentive compatible mechanism can only depend on ρ. This is not quite true as the mechanism could condition on η and c (in addition to ρ) whenever the agent is indifferent between multiple allocations. The result below establishes that, due to the linearity of the preference and the assumption that F admits a density, this can only happen on a set of types with zero probability.

Proposition 2 (Incentive Compatibility)**.** *A mechanism (x,t) is incentive compatible if and only if* *(i)* (x, t) *a.s. depends only on* $\rho = \theta/\eta$*;*

(ii) the allocation x *is non-decreasing in* ρ*; and*

(iii) the transfer as a function of the ratio ρ *satisfies*

$$t(\rho) = \rho v(x(\rho)) - \int_0^\rho v(x(r))dr + t(0)\,. \tag{5}$$

We thus write $x(\rho), t(\rho)$ and ignore types for which the allocation and transfer depend on more than the ratio ρ as these have probability 0 and thus play no role for the agents' or the designer's welfare. The ratio ρ is distributed according to the cumulative distribution function $H : \mathbb{R}_+ \to [0, 1]$ given by

$$H(z) = \mathbb{P}[\rho \le z] = \int \mathbb{1}_{\left\{\frac{\theta}{\eta} \le z\right\}} dF\,.$$

We denote the associated density by h:[22]

$$h(z) = \int \frac{\theta}{z^2} f\left(\theta, \frac{\theta}{z}\right) d\theta\,.$$

Using the result of Proposition 2, the following observation concludes that the self-financing constraint always binds in any optimal mechanism.

OBSERVATION 2. *There exists an optimal mechanism for the planner such that the budget feasibility constraint* (B) *binds, i.e.,*

$$\int t(\rho)dH = \kappa \int x(\rho)dH\,.$$

Further, (B) *binds for any optimal mechanism whenever* $\alpha < 1$.

Observation 2 is intuitive: the designer can always redistribute all potential surplus produced by the mechanism as an equal lump-sum payment to all agents without affecting incentive compatibility. If $\alpha = 1$ this change does not influence his value. Instead, if $\alpha < 1$, the planner cares about the agents' welfare and thus strictly benefits from redistributing any potential surplus. We henceforth without loss restrict attention to mechanisms where the budget constraint (B) binds.

Our next result characterizes the welfare of the planner only in terms of the allocation x given the transfer implied by (5). This result is the analog of Myerson's celebrated "Revenue equivalence" (Myerson, 1981).

[22] To see that this is indeed the correct density note that $H(z) = \int(1 - F_\eta(\theta/z|\theta))f_\theta(\theta)d\theta$, where $F_\eta(\cdot|\theta)$ is the CDF of η conditional on θ. Taking the derivative yields that $h(z) = \int \frac{\theta}{z^2} f_\eta(\theta/z|\theta)f_\theta(\theta)d\theta = \int \frac{\theta}{z^2} f(\theta, \theta/z)d\theta$.

PROPOSITION 3. *The welfare of the designer in a budget-balanced, incentive compatible mechanism (x, t) is given by*

$$\int (J(\rho)v(x(\rho)) - \kappa x(\rho))\, dH - \mathbb{E}\,[c]\, d(\bar{x}) - \kappa \bar{x}, \tag{6}$$

where the "virtual value" $J : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$J(z) = \alpha\, \mathbb{E}\,[\theta | \rho = z] + (1-\alpha) \left(z - (1 - \mathbb{E}\,[\eta \mid \rho \geq z]) \frac{1 - H(z)}{h(z)} \right). \tag{7}$$

Proposition 3 implies that the designer behaves as if they maximize allocative utility—but distorted, i.e. the designer assigns value $J(\rho)$ to agents with willingness to pay ρ. The designer thus treats rich agents with a low value for money and a low value for consuming the good the same as poor agents with a twice as high value for money who benefit twice as much from consuming the good. While the designer would benefit from treating these agents differently, they are not able to distinguish them in an incentive compatible way as we established in Proposition 2.

To understand the "virtual value" the designer maximizes, it is helpful to decompose it into three parts

$$J(z) = \alpha \underbrace{\mathbb{E}\,[\theta | \rho = z]}_{A} + (1-\alpha) \Big(\underbrace{z - \frac{1 - H(z)}{h(z)}}_{B} + \underbrace{\mathbb{E}\,[\eta | \rho \geq z]\, \frac{1 - H(z)}{h(z)}}_{C} \Big).$$

Term A represents the average value an agent with willingness to pay z derives from consuming the good, i.e., the expected allocative value of assigning the good to this agent. If $\alpha = 1$ and the designer is solely concerned with narrow allocative efficiency, this is the only term entering their virtual value.

Term B is the standard Myersonian virtual value derived from the willingness to pay. It measures the revenue raised by the mechanism. As all revenue raised by the mechanism is distributed back to the agents as a lump-sum payment and the average value assigned to money in the population equals 1, this term directly benefits the designer whenever $\alpha < 1$ and she cares about the agents' welfare.

The final term, C, represents the additional utility in information rents that accrue when an agent with willingness to pay z receives the good. More precisely, $\frac{1-H(z)}{h(z)}$ measures the additional information rent (measured in monetary terms) agents with willingness to pay ρ *larger* than z receive when an agent with willingness to pay z consumes the good. This term is multiplied by the average value for money, $\mathbb{E}\,[\eta | \rho \geq z]$, of agents with willingness to pay greater than z.

By a slight abuse of notation, let f, F denote the marginal distribution on (θ, η). The conditional expectation of the value for the good and money conditional on a given willingness to pay are then given by:[23]

$$\mathbb{E}\left[\theta | \rho = z\right] = \frac{\int \frac{\theta^2}{z^2} f(\theta, \theta/z) d\theta}{h(z)} = \frac{\int \theta^2 f(\theta, \theta/z) d\theta}{\int \theta f(\theta, \theta/z) d\theta}, \tag{8}$$

$$\mathbb{E}\left[\eta | \rho \geq z\right] = \frac{\int_z^\infty \int \frac{\theta^2}{r^3} f(\theta, \theta/r) d\theta dr}{1 - H(z)}. \tag{9}$$

The following properties of the virtual value will be useful to understand the incentives of the designer.

LEMMA 1. *The virtual value* $J : [0, \overline{\theta}/\underline{\eta}] \to \mathbb{R}$ *has the following properties:*

(i) J *is continuous,*
(ii) $J(0) = 0$*,*
(iii) $J(\overline{\theta}/\underline{\eta}) = \overline{\theta}(\alpha + (1-\alpha)1/\underline{\eta}) \geq \overline{\theta}$*.*

Given this representation of the designer's value, we can characterize the optimal mechanism. Before we can do this, we need one final definition. If J is not non-decreasing, let $\bar{J}$ be the usual ironed version of J: formally, for $q \in [0, 1]$,

$$\Psi(q) = \int_0^q J(H^{-1}(k)) dk.$$

Let $conv\,\Psi$ be the largest convex function that is pointwise smaller than Ψ. Convexity implies $conv\,\Psi$ differentiable almost everywhere. Define $\bar{J}(\rho) = \frac{\mathrm{d}}{\mathrm{d}q} conv\,\Psi(q)$ evaluated at $q = H(\rho)$ whenever the derivative of $conv\,\Psi$ is defined at q, and extend this to the entire interval by right continuity.

THEOREM 1. *The optimal mechanism* (x, t) *satisfies:*

$$x(\rho) \in \operatorname*{argmax}_{x \in X} \bar{J}(\rho) v(x) - (\mathbb{E}\left[c\right] d'(\bar{x}) + \kappa) x\,. \tag{10}$$

With transfer equal to

$$t(\rho) - t(0) = \rho v(x(\rho)) - \int_0^\rho v(x(r)) dr, \tag{11}$$

and the redistributive transfer $t(0)$ *is given by*

$$t(0) = -\int_0^{\overline{\theta}/\underline{\eta}} \left(\left[\rho - \frac{1 - H(\rho)}{h(\rho)} \right] v(x(\rho)) - \kappa x(\rho) \right) dH(\rho)\,.$$

[23]To see this, note that the joint density of (θ, ρ) is given by $f(\theta, \theta/\rho) \left| \frac{d\rho}{d\eta} \right| = f(\theta, \theta/\rho) \frac{\theta}{\eta^2}$. The density of θ conditional on ρ is thus given by $\frac{f(\theta, \theta/\rho) \frac{\theta}{\eta^2}}{h(\rho)}$.

Let $(v')^{-1}$ be the generalized inverse of v' and note that $(v')^{-1}(s) = 0$ for $s \geq 1$ or $s < 0$.[24] The consumption of an agent with willingness to pay ρ is given by

$$x(\rho) = v'^{-1}\left(\frac{\mathbb{E}\left[c\right] d'(\bar{x}) + \kappa}{\bar{J}(\rho)}\right). \tag{12}$$

Observe that Theorem 1 therefore clarifies how the optimal mechanism for the planner results in a consumption profile of the good that is different relative to our benchmarks of Laissez Faire (Definition 1, $x(\rho) = v'^{-1}(\kappa/\rho)$), Pigouvian Taxation (Definition 3, $x(\rho) = v'^{-1}\left((\mathbb{E}\left[c/\eta\right] d'(\bar{x}) + \kappa)/\rho\right)$), or Prohibition (Definition 2, $x \equiv 0$).

The exact nature of the distortion is not apriori obvious and depends on the details of the joint distribution of value for the good θ and value for money η in society. In Section 6, we illustrate this for two analytically tractable cases. In the rest of this section, we provide some general results. In a linear environment, we have the following simple corollary.

Corollary 2 (Linear Environments)**.** *Suppose $X = [0,1]$ and $v(x) = d(x) = x$.*

(i) It is optimal to prohibit consumption of the good iff $\bar{J}(\rho) \leq \mathbb{E}\left[c\right] + \kappa$ a.s..

(ii) Otherwise, charging a constant price per unit of consumption with full redistribution is optimal, and the optimal price ρ^* solves $\bar{J}(\rho^*) = \mathbb{E}\left[c\right] + \kappa$.

Remark 1. *Theorem 1 and Corollary 2 clarify a feature of the optimal mechanism that is not immediately obvious: the joint distribution of the externality cost c and the value for money η does not affect the optimal mechanism for the planner regardless of their objectives (allocative or utilitarian), only the expected externality $\mathbb{E}\left[c\right]$ matters. This is in contrast to the Pigouvian tax, which depends on the joint distribution of c and η (Definition 3). Intuitively, the optimal mechanism cares only about the total damage to society while Pigouvian taxes – which put higher weight on the rich – depend on whether the rich are more strongly affected by the externality. As remarked before, this feature of Pigouvian taxes seems normatively unappealing.*

As the next result shows, for general environments, one can implement the optimal allocation through a nonlinear consumption tax.

Proposition 4. *Suppose that $X = \mathbb{R}_+$ and v is strictly concave. The optimal allocation can be implemented by a tax that depends on consumption given by*[25]

$$t(x) - t(0) = (\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)) \int_0^x \frac{(x^*)^{-1}(q)}{\bar{J}((x^*)^{-1}(q))}\, dq, \text{ where,} \tag{13}$$

$$x^*(\rho) = v'^{-1}\left(\frac{\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)}{\bar{J}(\rho)}\right), \tag{14}$$

[24] That is $(v')^{-1}(s) = \sup\{x \in X \colon v'(x) \geq s\}$ with $\sup \emptyset = 0$.

[25] Here again, $(x^*)^{-1}$ denotes the pseudo inverse defined as $(x^*)^{-1}(q) = \sup\{\rho \colon x^*(\rho) \leq q\}$.

$$\bar{x}^* = \int x^*(\rho)dH,$$

It should be intuitive from the formulas that the implementing tax will generally be nonlinear. Formally, we have the following corollary:

COROLLARY 3. *Suppose that $X = \mathbb{R}_+$ and v is strictly concave and differentiable. If $\bar{J}(z)/z$ is increasing (decreasing), then the marginal tax on consumption is decreasing (respectively, increasing). A linear tax (constant marginal tax) on consumption is optimal if and only if J is linear.*

Corollary 3 straightforwardly suggests the following difference between mechanisms that are optimal for a planner with allocative/ utilitarian objectives and Pareto Efficiency:

REMARK 2. *The optimal mechanism under the allocative objective, utilitarian objective or any convex combination is generally* not *Pareto efficient, nor even aggregate-consumption-constrained Pareto efficient.*

To see this, note that Corollary 3 shows that the optimal tax is in general a nonlinear function of consumption, while we established that a mechanism is (aggregate-consumption-constrained) Pareto efficient if and only if it is a linear tariff (Corollary 1). While it is not surprising that the allocatively efficient mechanism is not (aggregate-consumption-constrained) Pareto efficient as the planner's preference, in this case, is not directly connected to the agents' preferences. However, it might be surprising that the optimal mechanism for a utilitarian planner is Pareto inefficient.

5.1. *Optimal Distortions Relative to Pigouvian Taxation*

This leads to the question of how and why the optimal mechanism deviates from Pigouvian taxes. As noted above, the utilitarian/ allocative optimal tax deviates from Pigouvian taxes by introducing nonlinearities. We next clarify the shape of these nonlinearities.

Utilitarian Objective:

PROPOSITION 5. *Consider a utilitarian planner. Assume $d(x) = x$ and J non-decreasing when J is positive (so that no ironing is necessary).*

(i) If $\lim_{\rho\to\infty}(1 - H(\rho))/h(\rho) = 0$ and $Cov(c, 1/\eta) \geq 0$, then there exists ρ' such that the consumption of agents with willingness to pay $\rho \geq \rho'$ is strictly higher in the optimal mechanism than under Pigouvian taxes.

(ii) Let $\mathbb{E}[\eta|\rho = 0] > 2$. For $\kappa, \mathbb{E}[c/\eta]$ low enough, there exists ρ'' that the consumption of agents with willingness to pay $\rho \leq \rho''$ is weakly lower in the optimal mechanism than under Pigouvian taxes and strictly lower for some ρ.

The above result identifies conditions such that, relative to the Pigouvian tax, the optimal tax imposed by a utilitarian planner on agents with a low willingness to pay is higher, and the optimal tax imposed on agents with a high willingness to pay is lower. The logic behind this proposition is as follows: Recall that when the planner has a utilitarian objective, the virtual value (7) reduces to:

$$J(z) = z - (1 - \mathbb{E}[\eta \mid \rho \geq z]) \frac{1 - H(z)}{h(z)}.$$

As we described earlier, this can be broken into two parts: $z - \frac{1-H(z)}{h(z)}$ describes the revenue contribution allocating the agent: the utilitarian planner here is concerned with revenue since all revenues are equally redistributed to the agents (who have an average value for money of 1). The $\mathbb{E}[\eta \mid \rho \geq z] \frac{1-H(z)}{h(z)}$ is the information rent to higher types from allocating a lower type.

Under the assumptions of part (i) of the proposition, the former term is close to z when z is large while the latter is positive. Therefore the utilitarian planner assigns the social value to high WTP agents that is not much smaller than their WTP. Further, the planner assigns the marginal externality cost of $\mathbb{E}[c]$ to consumption, while the Pigouvian tax considers the *monetary* externality cost of $\mathbb{E}[c/\eta]$. The result follows as we considered a good whose externality affects the rich weakly more, implying that $\mathbb{E}[c/\eta] > \mathbb{E}[c]$. Therefore, a utilitarian planner will allow for more consumption of agents with a high WTP than under Pigouvian taxation: this planner values the additional revenue they generate (and redistribute) by such allocation. For goods $\mathrm{Cov}(c, 1/\eta) < 0$, i.e. whose consumption affects the poor more strongly, this result need not hold. In this case, a Pigouvian tax will be lower as it is cheaper to compensate the poor for the externality, which can decrease Pigouvian taxes below the highest marginal tax of a utilitarian planner.

Part (ii) is more subtle: If $\mathbb{E}[\eta|\rho = 0] > 2$ then agents whose WTP equals zero value money twice as much as the average in the population $\mathbb{E}[\eta] = 1$, i.e. are relatively poor. We establish that this implies virtual value is negative for an interval around zero. Intuitively, allocating the good to agents with a low WTP creates minor allocative benefits, but substantially depresses the revenue raised by the mechanism as it increases the information rents paid to higher types.[26] As the utilitarian planner values redistribution, he would rather not allocate to these low WTP types. In contrast, a Pigovuian planner will allocate to every agent whose WTP exceeds the externality cost, which by assumption of the proposition is not too high.

[26]The exact same condition that the benefit of redistributing a dollar to the poorest agent exceeds the social value of a dollar appears in various papers Kang (2022); Kang and Zheng (2020); Akbarpour et al. (2020). This is no coincidence and Dworczak et al. (2022) shows why this condition naturally appears in allocation problems where agents value money differently when the distribution of values for money admits a density.

Thus, maybe surprisingly, a utilitarian planner (who cares more about the poor than a planner for whom a Pigouvian tax is optimal) will reduce the consumption of agents with a low WTP (who in the independent case are more likely to be poor) and increase the consumption of those with a high WTP (who in the independent case are more likely to be rich). Changing the consumption profile this way increases revenue, and as all taxes are redistributed, benefits the poor. For example, a fuel tax above the level implied by the monetary externality can be justified by its additional redistributive benefit.

Allocative Efficiency Objective: The following proposition summarizes the shape of the distortions relative to a Pigouvian (linear) tax for a planner concerned with narrow allocative efficiency.

PROPOSITION 6. *Consider a planner concerned with narrow allocative efficiency. Assume that* $d(x) = x$*, J is non-decreasing (so that no ironing is necessary), and* $\mathbb{E}[\eta|\rho = z]$ *is decreasing in z. Then, in the optimal mechanism for an allocative efficient planner:*

(i) If $\lim_{z\to\infty} \mathbb{E}[\eta|\rho = z] \to 0$*, there exists* $\bar{\rho}$ *such that the consumption of agents with WTP* $\rho \geq \bar{\rho}$ *is strictly lower than under Pigouvian taxes.*

(ii) If $Cov(c, 1/\eta) \geq 0$ *and* $\kappa, \mathbb{E}[c]$ *are low enough, there exists* $\underline{\rho}$ *such that the consumption of agents with WTP to pay* $\rho < \underline{\rho}$ *is weakly higher than under Pigouvian taxes.*

The logic behind this proposition is as follows: when the planner has a narrow allocative efficiency objective, the virtual value (7) reduces to:

$$J(z) = \mathbb{E}[\theta|\rho = z] = z\,\mathbb{E}[\eta|\rho = z] \ .$$

Part (i) follows because under the maintained assumptions, for z large, $J(z) < z$ and therefore the allocative efficient planner would like to allocate the agents with high WTP less than they would under Pigouvian taxation. Conversely, under the maintained assumptions, since $\mathbb{E}[\eta] = 1$, it must be that $\mathbb{E}[\eta|\rho = z] \geq 1$ for z small. Further, since $\text{Cov}(c, 1/\eta) \geq 0$, we have that $\mathbb{E}[c/\eta] > \mathbb{E}[c]$. Therefore if $\kappa, \mathbb{E}[c]$ are small enough, there exists $\underline{\rho} > \kappa + \mathbb{E}[c]$ such that the allocative efficient planner would like to allocate agents with WTP smaller than $\underline{\rho}$ more than under the Pigouvian tax.

Proposition 5 and 6 thus show that the qualitative nature of the optimal tax depends on the planner's objective. Relative to the Pigouvian benchmark, a utilitarian planner will impose higher taxes on agents with a low WTP (who are more likely to be poor) and lower taxes on agents with a high WTP (who are more likely to be rich). A planner interested in allocative efficiency will distort taxes in the opposite way.

6. EXAMPLES

We illustrate our characterization of optimal mechanisms using two especially tractable distributions. First, we consider the case where both value for the good and value for money are distributed independently according to the Rayleigh distribution. Next, we consider the case where the two are jointly lognormally distributed (and potentially correlated). For simplicity, we assume that the externality disutility agents suffer is linear in the total consumption, i.e., $d(\bar{x}) = \bar{x}$, and in addition that the extent to which people care about the externality c is distributed independently of their value for the good and money. We obtain analytical closed forms for the virtual value of the agent as a function of their willingness to pay, ρ, which allows us to explicitly characterize the optimal mechanism.

6.1. *Rayleigh distribution*

Our first setting is when both the agent's value for the good θ and the value for money η are independently distributed according to the Rayleigh distribution. Recall that the Rayleigh distribution with parameter σ^2 has PDF $f(x) = \frac{x}{\sigma^2}\exp(-\frac{x^2}{2\sigma^2})$, and CDF $F(x) = 1 - \exp(-\frac{x^2}{2\sigma^2})$. In particular, the former is distributed with parameter σ_θ^2, while the latter is distributed with parameter $\sigma_\eta^2 = 2/\pi$.[27]

PROPOSITION 7. *Suppose the agent's value for the good and value for money are both distributed according to the Rayleigh distribution as described above. Then, we have that*

$$J(z) = \alpha \frac{3z\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}} + (1-\alpha)\left(z - \left(1 - \frac{\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}}\right)\frac{z^2\sigma_\eta^2 + \sigma_\theta^2}{2z\sigma_\eta^2}\right). \tag{15}$$

No ironing is needed in the optimal mechanism, i.e., $\overline{J}(z) = J(z)$. Further, if v is strictly concave, the tax on consumption that implements the optimal mechanism is such that marginal taxes are increasing in consumption, for either objective of the planer.

The analytical form allows us to summarize these distortions in a straightforward corollary. In particular, note that the comparison of the allocative efficient mechanism to the Pigouvian tax is as suggested by Proposition 6.

COROLLARY 4. *Suppose (θ, η) are distributed according to the Rayleigh distribution.*

(i) For a planner with the allocative objective ($\alpha = 1$), the optimal mechanism features a consumption cap, i.e., a maximum amount that is allowed to be consumed by any agent. If $\mathbb{E}[c]$ is small enough, some agent types consume more *than they would under Laissez-faire.*

(ii) For a planner with the utilitarian objective ($\alpha = 0$), the optimal mechanism has no consumption cap.

[27] This choice of σ_η^2 ensures that the average value for money, $\mathbb{E}[\eta]$ equals 1 as required by our normalization.

It is helpful to discuss briefly the intuition for this corollary. First, consider a planner concerned with narrow allocative efficiency, i.e., $\alpha = 1$: In this case, recall, we have that $J(z) = \mathbb{E}[\theta|\rho = z]$. The specific functional form from (15) tells us that for "low" WTP agents, the planner would actually allow for more consumption than they would have consumed under the Laissez-faire or the Pigouvian tax mechanisms. This is because, under this distribution, the planner views low willingness to pay as relatively likely to result from high values for money, so that $J(\rho) > \rho$. Therefore relative to either mechanism, the optimal mechanism for the planner will involve more consumption on both the intensive and extensive margins. Further, it is easy to see that $J(\rho)$ asymptotes to $3\frac{\sigma_\theta}{\sigma_\eta}$, so that the optimal mechanism in this case for a planner features a "cap" on the maximum amount any agent can consume. A related result is obtained in Proposition 5 of Kang (2022), for the case where agents may differ in the externality they cause and a different allocative objective of the planner. In contrast to the consumption cap that is optimal for the allocative objective, under both Laissez-Faire and Pigouvian taxation, there is no bound on the consumption of the agents.

Next, consider a planner concerned with utilitarian efficiency, i.e., $\alpha = 0$: In this case, note that as ρ grows large, $J(\rho) \approx \rho/2$, and thus consumption grows unboundedly as willingness to pay grows. This is because even though high consumption by these types is not allocatively efficient (recall the discussion in the former case). The reason that the utilitarian planner nevertheless does not want to limit consumption is that the gains from taxing high types for their consumption and redistributing the proceeds exceed the allocative inefficiency.

In the linear case, the optimal mechanism is a constant tax per unit (Corollary 2) and we are able rank optimal tax under narrow allocative and utilitarian objectives vis-a-vis the Pigouvian price. The ranking of optimal taxes reverses depending on whether externalities are high or low.

COROLLARY 5. *Consider (θ, η) distributed according to the Rayleigh distribution, $v(x) = x$ and $\kappa = 0$. Let p_P be the Pigouvian tax, p_A the allocative efficient price and p_U the utilitarian efficient price. There exist two cutoffs $\underline{c}$ and $\overline{c}$ such that:*

(i) For $\mathbb{E}[c] < \underline{c}$ we have $p_A < p_U < p_P$.

(ii) For $\mathbb{E}[c] > \overline{c}$ we have $p_P < p_U < p_A$.

6.2. *Lognormal Distribution*

Suppose that the value for the good and for money are jointly lognormally distributed with the cost independent of θ and η, i.e.,

$$\begin{pmatrix} \log\theta \\ \log\eta \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_\theta \\ -\sigma_\eta^2/2 \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \xi\sigma_\theta\sigma_\eta \\ \xi\sigma_\theta\sigma_\eta & \sigma_\eta^2 \end{pmatrix} \right), \tag{16}$$

where $\xi \in [-1, 1]$ is the correlation between $\log\theta$ and $\log\eta$. Assuming that $\log\eta \sim \mathcal{N}(-\sigma_\eta^2/2, \sigma_\eta^2)$ implies that $\mathbb{E}[\eta] = 1$.

REMARK 3. *If value for the good and value for money are jointly lognormally distributed as above, then they are jointly lognormally distributed with $\rho = \frac{\theta}{\eta}$ and*

$$\begin{pmatrix} \log\theta \\ \log\eta \\ \log\rho \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_\theta \\ -\sigma_\eta^2/2 \\ \mu_\rho \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \xi\sigma_\theta\sigma_\eta & \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta \\ \xi\sigma_\theta\sigma_\eta & \sigma_\eta^2 & \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2 \\ \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta & \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2 & \sigma_\rho^2 \end{pmatrix} \right),$$

where $\mu_\rho = \mu_\theta + \sigma_\eta^2/2$ and $\sigma_\rho^2 = \sigma_\theta^2 + \sigma_\eta^2 - 2\xi\sigma_\theta\sigma_\eta$.

Remark 3 follows from standard properties of the lognormal distribution, and we provide a proof in the Appendix. Given that the joint distribution of (θ, η, ρ) is lognormal, we can compute $\mathbb{E}[\theta|\rho = z]$ and $\mathbb{E}[\eta|\rho \geq z]$ in closed form as[28]

$$\mathbb{E}[\theta|\rho = z] = \mathbb{E}[\theta] \frac{z^\beta}{\mathbb{E}[\rho^\beta]} \qquad \mathbb{E}[\eta|\rho \geq z] = \frac{1 - H\left(z e^{(1-\beta)\sigma_\rho^2}\right)}{1 - H(z)}.$$

The constant β measures how good of a signal willingness-to-pay is as a signal about value and is given by

$$\beta = \frac{\mathrm{Cov}(\theta, \rho)}{\sigma_\rho^2} = \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\theta^2 + \sigma_\eta^2 - 2\xi\sigma_\theta\sigma_\eta} = \frac{(\sigma_\theta/\sigma_\eta) - \xi}{(\sigma_\eta/\sigma_\theta) + (\sigma_\theta/\sigma_\eta) - 2\xi}. \tag{17}$$

The informativeness β is decreasing in the ratio of the variance in the value for money σ_η^2 relative to the value for the good σ_θ^2, and vanishes, i.e $\beta \to 0$, if the ratio $\sigma_\eta^2/\sigma_\theta^2$ goes to infinity. Intuitively, there are two reasons an agent has a high WTP: (i) they might either value the good a lot, or (ii) might value money very little. If the variation in values for money is relatively large the latter becomes a more likely explanation and the WTP a worse signal for an agent's value. Note that $\beta < 1$ if the dispersion in values for money exceeds the dispersion in values for the good, i.e. $\sigma_\eta > \sigma_\theta$.[29] We can now describe the virtual value and characterize the optimal mechanism in this setting.

[28]We establish these statements as part of the proof of Proposition 8 in the Appendix. The expectation $\mathbb{E}[\rho^\beta]$ equals $\exp\left(\beta\mu_\rho + \beta^2 \frac{1}{2}\sigma_\rho^2\right)$.

[29]This is plausibly the typical case given the common large variation in wealth in many societies.

PROPOSITION 8. *Suppose the agent's value for the good and value for money are jointly lognormally distributed as in* (16)*. Then, the virtual value J is given by*

$$J(z) = \alpha\,\mathbb{E}\left[\theta\right]\frac{z^{\beta}}{\mathbb{E}\left[\rho^{\beta}\right]} + (1-\alpha)\left(z - \frac{H\left(ze^{(1-\beta)\sigma_\rho^2}\right) - H(z)}{h(z)}\right). \tag{18}$$

Here H, h are the CDF, PDF of ρ, and $\beta = Cov(\theta,\rho)/\sigma_\rho^2$. The "ironed virtual value" under the allocative objective $\alpha = 1$ and the utilitarian objective $\alpha = 0$ equals

$$\bar{J}(z) = \begin{cases} \mathbf{1}_{\beta\geq 0}J(z) + \mathbf{1}_{\beta<0}\mathbb{E}\left[\theta\right] & \text{for } \alpha = 1 \\ \max\left\{0, J(z)\right\} & \text{for } \alpha = 0\,. \end{cases}$$

Given the above characterization of the virtual value we can characterize the optimal mechanism as follows:

PROPOSITION 9 (Optimal Mechanism in the Lognormal Case)**.** *Suppose (θ,η) are jointly lognormally distributed and v is strictly concave. Consider a planner concerned with narrow allocative efficiency, i.e., $\alpha = 1$.*

(i) $\beta < 0$: there exists a threshold $\bar{c} > \mathbb{E}\left[\theta\right]$ such that if $\mathbb{E}\left[c\right] \leq \bar{c}$, the optimal mechanism is a consumption quota and if $\mathbb{E}\left[c\right] > \bar{c}$, the optimal mechanism is prohibition.

(ii) $\beta > 0$: the optimal mechanism is a consumption tax with strictly increasing marginal tax if $\beta \in (0,1)$ and strictly decreasing marginal tax if $\beta > 1$.

Consider a utilitarian planner, i.e., $\alpha = 0$.

(iii) The optimal mechanism is a consumption tax with strictly decreasing marginal tax.

It is worth emphasizing the result for a utilitarian objective: marginal taxes in this case are *decreasing* in consumption. This is despite a natural objective function of the planner (utilitarian) and, arguably, a standard assumption distribution, i.e., lognormal. This illustrates, more explicitly, the distortions relative to the Pigouvian tax described in Proposition 5— and the intuition is similar. High WTP types will consume more under a decreasing marginal tax than under Pigouvian taxation. The utilitarian objective implies that the principal values this consumption and resulting taxation since they can redistribute the proceeds. To our knowledge, diminishing marginal taxes have not been discussed in the public discourse, even though this discourse may implicitly consider a utilitarian objective for the planner (since it values redistribution).

The optimal mechanism under the allocative objective takes a non-market form for $\beta < 0$, or equivalently $\sigma_\theta/\sigma_\eta < \text{cor}(\theta,\eta) = \xi$. This is thus only the case for goods that are liked more by poor people (who have a higher value for money). For such goods, a high willingness to pay is a negative signal about the agent's value for the good. A planner with an allocative objective would hence prefer to allocate more of the good to agents with a *lower* willingness to pay. Of course, this is not incentive compatible. Therefore, the optimal allocation is a constant quota for all agents, which can be financed through a uniform tax if the externality cost is not too high.[30] If the externality cost is high enough, in particular, if the average value for the good in society is lower than the cost of the externality, then prohibition is optimal.

The following corollary shows that it is possible that the optimal mechanism may involve *more* consumption than even Laissez-Faire (Definition 1):

COROLLARY 6. *Suppose (θ, η) are jointly lognormally distributed and $\beta > 1$. Then for a utilitarian planner, i.e. $\alpha = 0$,*

(i) For $\mathbb{E}[c]$ small enough, the optimal mechanism has uniformly higher consumption than under Laissez-faire.
(ii) For β close to 1, the optimal mechanism will have lower consumption than under Laissez-Faire but higher consumption than under Pigouvian taxation.

Note that in order to incentivize higher consumption, the planner will have to tax agents who do not consume the good (low willingness to pay). It is easy to see that $\beta > 1$ whenever $\xi > \sigma_\eta / \sigma_\theta$, i.e. the correlation between value for money and value for the good is relatively large and in addition the values for the good are more dispersed than the values for money $\sigma_\theta > \sigma_\eta$. In such a case, low WTP agents are those with low values for the good and low value for money, and therefore a utilitarian objective planner is better off taxing such agents and subsidizing the consumption of high WTP agents (who have high value for the good/ high value for money). This can be achieved by a lump sum tax on all agents, which is rebated as a per-unit subsidy on the good.

7. ENDOGENOUS WELFARE WEIGHTS

Up to now, we assumed that the value a utilitarian planner assigns to transferring money from one agent to another is exogenously given, and the agents' preferences over money are linear. We next relax these assumptions and illustrate how these "welfare weights" used by the planner can be determined endogenously for a utilitarian planner. More formally, we consider agents with different income levels i and mechanisms where the designer cannot directly condition on the income level. Each agent is characterized by their willingness to pay for consuming the good ρ_θ, their willingness to pay for reducing aggregate consumption ρ_c, and their income i. An agent with type (ρ_θ, i, ρ_c) who consumes x and makes a transfer of t to the mechanism given total consumption $\bar{x}$ in society has a utility of

$$\varphi\left(\rho_\theta v(x) + i - t - \rho_c \bar{x}\right) . \quad (19)$$

[30] A real world example of such a fixed consumption quota being made available to all agents for free can be found in the case of public schooling (of course, this is a positive externality).

Here, $\varphi : \mathbb{R} \to \mathbb{R}$ is a strictly concave and strictly increasing function. As the agent's preferences are unaffected by the monotone transformation φ, Proposition 2 provides a characterization of incentive compatible mechanisms. We can define

$$U(\rho_\theta, i, \rho_c; x, t) = \rho_\theta v(x(\rho_\theta, \rho_c)) + i - t(\rho_\theta, \rho_c) - \rho_c \bar{x},$$

and write the problem of a utilitarian designer as

$$\max_{(x,t)} \int \varphi\left(U(\rho_\theta, i, \rho_c; x, t)\right) dF\,. \tag{20}$$

An additional dollar to an agent of type (ρ_θ, i, ρ_c) increases the designer's utility by

$$\varphi'(U(\rho_\theta, i, \rho_c; x, t)).$$

A natural conjecture, therefore, is that the optimal mechanism in this setting can be obtained as the mechanism that is optimal in our original quasilinear setting with associated *endogenously derived* values for money $\eta = \varphi'(U(\rho_\theta, i, \rho_c; x, t))$. Our next proposition shows that this is indeed the case.

PROPOSITION 10. *A mechanism (x, t) maximizes (20) subject to (B) and (IC) if and only if it maximizes (OBJ) for $\alpha = 0$ subject to (B) and (IC) with respect to the type distribution F obtained by setting*

$$\begin{aligned} \theta &= \rho_\theta\, w(\rho_\theta, i, \rho_c) \\ \eta &= w(\rho_\theta, i, \rho_c) \\ c &= \rho_c\, w(\rho_\theta, i, \rho_c) \end{aligned}$$

where $w(\rho_\theta, i, \rho_c) = \varphi'(U(\rho_\theta, i, \rho_c; x, t))$.

The above result thus establishes that solving for the optimal mechanism when the agent's utilities are nonlinear and given by (19) is equivalent to a fixed point problem where the values for money depend on the mechanism and are endogenously determined. Our previous characterization of optimal utilitarian mechanisms for the quasilinear case, thus completely translates to the more general case presented in this section. As φ' is decreasing, richer agents will be endogenously assigned a lower value for money by the planner. If the part of the agents' utility determined by the mechanism is small relative to the agents' overall utility, assuming w to be exogenous is a good approximation.

8. AN ILLUSTRATION

We next illustrate our results using the concrete example of air travel. Air travel is a prime example of an externality-causing activity: it is estimated that it is responsible

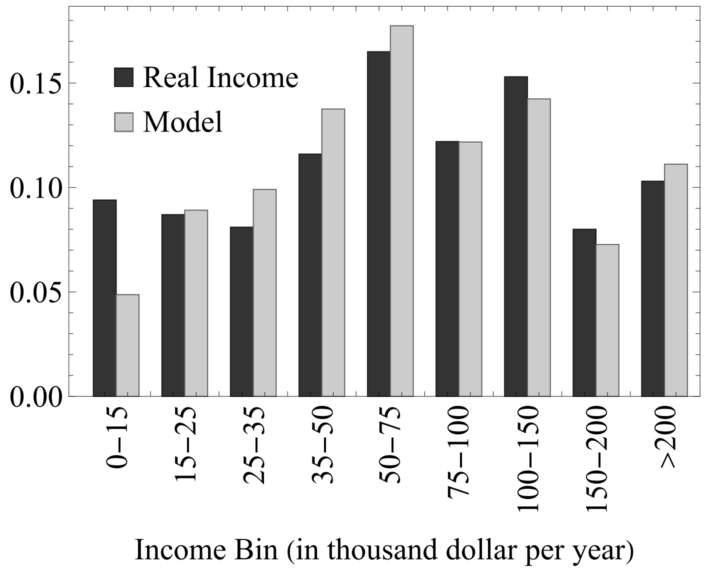

FIGURE 1. Real and calibrated household income distribution.

for 5% of global warming and 2.4% of CO2 emissions (see Timperley, 2020). Furthermore, given that individuals are already required to provide detailed personal information when booking flights, and tickets cannot be resold, implementing a non-linear tax based on (say) annual air-travel seems quite feasible. Air travel also only constitutes a small enough proportion of (most) peoples' annual incomes which makes our linearity assumptions reasonable.

We calibrate the special case of our model introduced in Section 6.2 where the joint distribution of agents' value for flying and value for money is lognormal. As is standard in the public finance literature, to analyze the utilitarian objective we have to make an assumption that pins down the planner's gains from redistribution as these can not be determined from individual preferences. For simplicity, we assume that the planner uses welfare weights of the form $i^{-\gamma}$, where i is household income and γ is chosen such that the planner is indifferent to taking \$2 from a household making \$100k annually and giving \$1 to one making \$10k. Finally, we also assume that household income is lognormally distributed (with $\mu = 4.2$ and $\sigma^2 = 0.81$) which, as we show in Figure 1, roughly matches the US household income distribution (US Census Bureau).[31]

We measure consumption of air travel in round trips and set the production cost κ equal to \$459 which is the inflation adjusted average price of a round trip originating from the US in 2015.[32] We assume that the marginal value of an additional trip is of power form,

[31]See page 16 of https://www.census.gov/content/dam/Census/library/publications/2022/demo/p60-276.pdf.

[32]See Bureau of Transportation Statistics (2022).

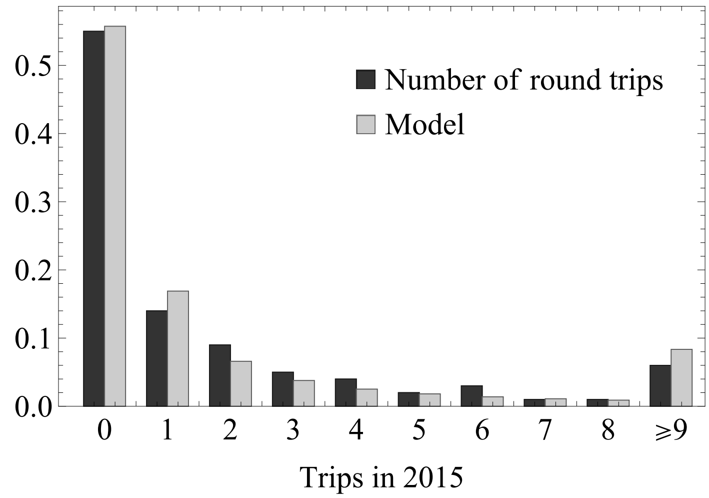

FIGURE 2. Real and calibrated number of round trips by household.

$v'(x) = x^{-\alpha}$ which is equivalent to assuming a constant elasticity of $-1/\alpha$.[33] We set the elasticity to -1.146 which is the mean elasticity found in the meta-study of price elasticities of demand for passenger air travel of Brons et al. (2002). We calibrate the distribution of willingness to pay to be lognormal with parameters $\mu_\rho = 5.2, \sigma_\rho^2 = 4.4$ based on data from the "Air Travelers in America: Annual Survey"[34] collected in 2015.[35] As one can see in Figure 2 this calibration fits the real distribution of flights well. To calibrate the correlation ξ between value for air-travel and value for money we use aggregate data on the flying habits of individuals in different wealth brackets that was collected in a YouGov survey in 2015.[36] The survey measures the number of flights taken by people with different household income and thus can be used to measure the correlation between WTP and income which then identifies $\xi = -3.3$. Figure 3 shows the flight data by income and the prediction of our calibrated model. This value is consistent with rich people deriving a higher utility from flying. Table 1 shows the optimal taxes under the utilitarian and allocative objective assuming that the Pigouvian tax equals \$50.[37] As one can see

[33]To see that this is the elasticity observe that the demand at price π equals $\mathbb{E}[x] = \mathbb{E}[(\rho/\pi)^{1/\alpha}]$ and $\partial\mathbb{E}[x]/\partial\pi = -\frac{1}{\alpha\pi}\mathbb{E}[(\rho/\pi)^{1/\alpha}]$ and hence $\pi \times (\partial\mathbb{E}[x]/\partial\pi)/\mathbb{E}[x] = -1/\alpha$. Note here that we did not normalize $v'(0) = 1$ as this is more convenient for the calibration of the model.

[34]The survey was run by Ipsos and includes roughly 10000 individuals see https://www.airlines.org/dataset/air-travelers-in-america-annual-survey/.

[35]The specific value we have chosen maximize the model fit in terms of squared distance.

[36]The survey data is available at https://www.statista.com/statistics/316376/air-travel-frequency-us-by-income/. We calibrate our model to the survey data by assuming that the anser people give in the survey is a monotone function of WTP, i.e. people with low WTP report "no air travel", people with medium WTP report "occasional air travel", and people with high WTP report "frequent air travel".

[37]The average length of a US round trip is around 500 miles (see https://www.statista.com/statistics/742763/regional-carriers-average-passenger-trip-length/). According to Atmos that produces roughly 0.5 tons of emissions per round trip (https://www.atmosfair.de/en/offset/flight/), so \$50

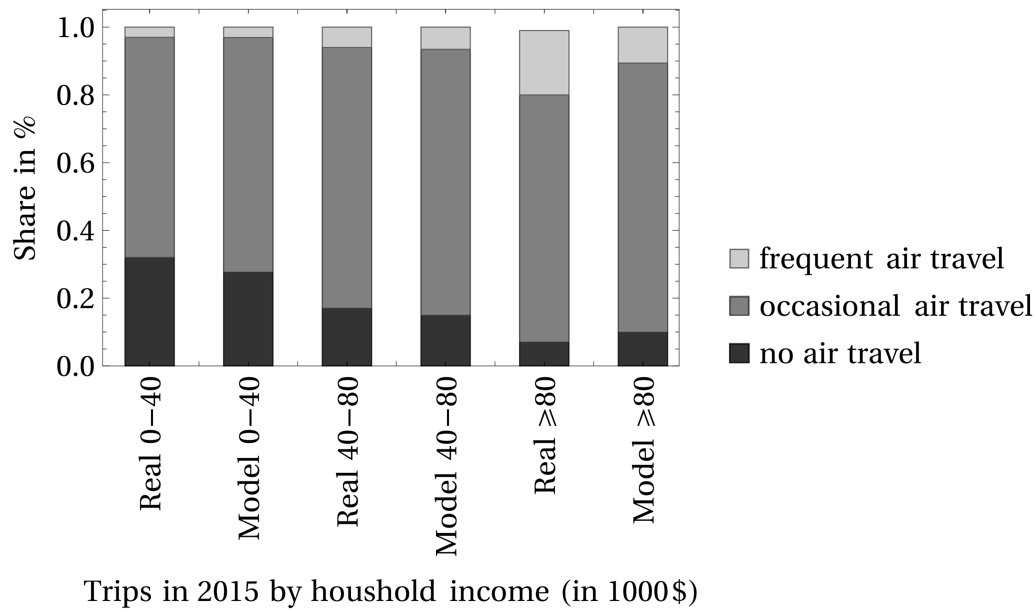

FIGURE 3. Real and calibrated flights by household income.

	Avg Trips	Rebate	Flight Taxes 1st	2nd	3rd	4th	5th	6th
Allocative	3.55	1037.11	124.58	168.78	186.55	198.38	207.31	214.52
Utilitarian	4.27	794.79	215.10	207.46	204.76	203.05	201.79	200.80
Pigouvian	5.63	281.5	50	50	50	50	50	50

TABLE 1. Table summarizing the average number of round trips, average redistribution to all agents and marginal tax of 1st thru 6th flights under Pigouvian, Allocative Efficient and Utilitarian Optimal Taxation.

optimal taxes under both objectives are much higher than the Pigouvian tax, between \$124-\$215. Intuitively, a utilitarian planner likes higher taxes as he values the additional benefit of redistribution and an allocative planner would like higher taxes as WTP due to differences in values for money is a poor signal about how much the planner values consumption of a given agent. Utilitarian taxes are roughly constant around \$200 per flight while allocative taxes are increasing in the number of flights. Both optimal taxes lead to a significant rebate to households of between \$795 and \$1037 per year. As 55% of the population do not fly in a given year it seems plausible that such a tax would make the majority of households better off while at the same time reducing emissions.

value is (roughly) in line with a \$100 per ton carbon price. In Table 2 we present the results after halving the carbon price to \$50 per ton. The taxes are qualitatively similar, albeit ≈20% lower.

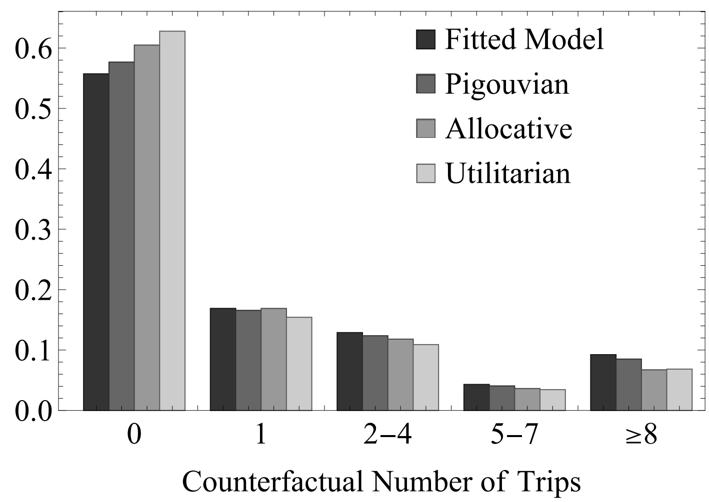

FIGURE 4. Distribution of the number of flights taken by individuals under Pigouvian, Allocative Efficient and Utilitarian Optimal Taxation.

9. CONCLUSIONS

This paper considered the taxation of an externality causing good or activity in an unequal society, i.e., one where agents have different values for money. While economists have often pointed to Pigouvian taxation as the main policy instrument to regulate such activity, this has been criticized in the public discourse as regressive. We showed within the context of the model that this is indeed the case: Pigouvian taxes correspond to the planner putting higher Pareto weights on richer agents, i.e., have lower values for money. In that sense, these critiques may indeed have merit.

We considered a rich class of objectives for the planner, encompassing any convex combination of allocative objective (maximizing the net value from consumption less the social cost, ignoring transfers) and utilitarian objective (maximizing the net utility in society). Our techniques can be straightforwardly adapted if the planner additionally has a revenue requirement that must be met by the raised taxes, or if the planner would like to apply some other welfare weights to the utilities.

Despite rich private information of the agents (value for the good, value for money and disutility from externalities), we fully characterize the optimal feasible mechanism for the planner. The way in which the optimal mechanism differs from Pigouvian taxes depends on the objective of the designer. A planner interested in utilitarian welfare will induce lower consumption of agents with a low WTP (more likely to be poor) and increase consumption of high WTP agents (more likely to be rich) as doing so increases the tax revenue and thus the amount of money available for redistribution. In contrast, a planner interested in narrow allocative efficiency will increase the consumption of low WTP agents and decrease the consumption of high WTP agents as a poor person's willingness to pay understates their benefit of consumption, and a rich person's willingness to pay overstates their benefit of consumption. In general, these mechanisms will be implemented by nonlinear taxes, while Pareto efficient mechanisms for any level of consumption in society correspond to linear taxes.

We explicitly derived the optimal tax for two joint distributions of value for the good and value for money—independent Rayleigh and joint log normal. In these cases, the optimal mechanism involves features proposed in public regulation: full prohibition of the activity, a consumption cap on the maximum amount any agent can consume, or a uniform quota. We then illustrated the applicability of the model by calibrating it US air-travel. Given the tractability of our framework, we hope that there are many more applications. Another context where we think quantity-based non-linear taxes can be implemented is distance based pricing for road vehicles: indeed, Singapore is already in the process of rolling out GPS-based tracking devices on all cars, making accurate measurement of distance traveled feasible.[38]

While we applied our framework to a situation where the agents are citizens and the designer decided on taxation within a country, it might be useful in other contexts. One example where this might be useful is to think about the design of climate treaties where rich and poor countries (can afford to) care differentially about the damage caused by carbon emissions.

References

Akbarpour, Mohammad, Eric B Budish, Piotr Dworczak, and Scott Duke Kominers, "An economic framework for vaccine prioritization," *Quarterly Journal of Economics*, 2022.

______, **Piotr Dworczak, and Scott Duke Kominers**, "Redistributive allocation mechanisms," *Available at SSRN 3609182*, 2020.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 2019, *134* (3), 1557–1626.

______, **Christopher Knittel, and Dmitry Taubinsky**, "Tagging and Targeting of Energy Efficiency Subsidies," *American Economic Review*, May 2015, *105* (5), 187–91.

Allen, Roy and John Rehbeck, "Obstacles to Redistribution Through Markets and One Solution," *Available at SSRN 3966532*, 2021.

Atkinson, Anthony Barnes and Joseph E Stiglitz, "The design of tax structure: direct versus indirect taxation," *Journal of public Economics*, 1976, *6* (1-2), 55–75.

[38]See, e.g., https://www.straitstimes.com/singapore/transport/new-erp-system-to-start-2023-but-no-distance-based-charging-just-yet.

Baumol, William J, "On taxation and the control of externalities," *The American Economic Review*, 1972, *62* (3), 307–322.

Borenstein, Severin and Lucas W Davis, "The distributional effects of US clean energy tax credits," *Tax Policy and the Economy*, 2016, *30* (1), 191–234.

Brons, Martijn, Eric Pels, Peter Nijkamp, and Piet Rietveld, "Price elasticities of demand for passenger air travel: a meta-analysis," *Journal of Air Transport Management*, 2002, *8* (3), 165–175.

Bureau of Transportation Statistics, https://www.bts.gov/content/annual-us-domestic-average-itinerary-fare-current-and-constant-dollars 2022.

Condorelli, Daniele, "Market and non-market mechanisms for the optimal allocation of scarce resources," *Games and Economic Behavior*, 2013, *82*, 582–591.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet, "Direct versus indirect taxation: the design of the tax structure revisited," *International Economic Review*, 2001, *42* (3), 781–800.

Davis, Lucas W and Christopher R Knittel, "Are fuel economy standards regressive?," *Journal of the Association of Environmental and Resource Economists*, 2019, *6* (S1), S37–S63.

Diamond, Peter A, "Consumption externalities and imperfect corrective pricing," *The Bell Journal of Economics and Management Science*, 1973, pp. 526–538.

Dixit, Avinash, "Tax policy in open economies," in "Handbook of public economics," Vol. 1, Elsevier, 1985, pp. 313–374.

Dorfman, Jeffrey, "Uncomfortable Truths About Climate Change, Carbon Taxes And Inequality," *Forbes*, 2018.

Dworczak, Piotr et al., "Equity-efficiency trade-off in quasi-linear environments," Technical Report, GRAPE Group for Research in Applied Economics 2022.

______, Scott Duke Kominers, and Mohammad Akbarpour, "Redistribution through markets," *Econometrica*, 2021, *89* (4), 1665–1698.

Ferey, Antoine, Benjamin B Lockwood, and Dmitry Taubinsky, "Sufficient Statistics for Nonlinear Tax Systems with General Across-Income Heterogeneity," 2022.

Fullerton, Don and Gilbert E Metcalf, "Environmental taxes and the double-dividend hypothesis: did you really expect something for nothing?," 1997.

Goldin, Jacob and Tatiana Homonoff, "Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity," *American Economic Journal: Economic Policy*, February 2013, *5* (1), 302–36.

Herweg, Fabian and Klaus M Schmidt, "How to Regulate Carbon Emissions with Climate-conscious Consumers," 2022.

Kang, Mingshi and Charles Z Zheng, "Pareto Optimality of Allocating the Bad," Technical Report, Mimeo, December 2020. 1, 1 2020.

Kang, Zi Yang, "Markets for goods with externalities," *Available at SSRN 3586050*, 2022.

Kopczuk, Wojciech, "A note on optimal taxation in the presence of externalities," *Economics Letters*, 2003, *80* (1), 81–86.

Milgrom, Paul and Ilya Segal, "Envelope theorems for arbitrary choice sets," *Econometrica*, 2002, *70* (2), 583–601.

Mirrlees, J.A., "Optimal tax theory: A synthesis," *Journal of Public Economics*, 1976, *6* (4), 327–358.

Myerson, Roger B, "Optimal auction design," *Mathematics of operations research*, 1981, *6* (1), 58–73.

Negishi, Takashi, "Welfare economics and existence of an equilibrium for a competitive economy," *Metroeconomica*, 1960, *12* (2-3), 92–97.

Pigou, AC, *The Economics of Welfare*, Palgrave Macmillan, 1920.

Rothschild, Casey and Florian Scheuer, "Optimal Taxation with Rent-Seeking," *The Review of Economic Studies*, 03 2016, *83* (3), 1225–1262.

Saez, Emmanuel, "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 2002, *83* (2), 217–230.

______ **and Stefanie Stantcheva**, "Generalized social marginal welfare weights for optimal tax theory," *American Economic Review*, 2016, *106* (1), 24–45.

Sallee, James M, "Pigou creates losers: On the implausibility of achieving pareto improvements from efficiency-enhancing policies," Technical Report, National Bureau of Economic Research 2019.

Sandmo, Agnar, "Optimal taxation in the presence of externalities," *The Swedish Journal of Economics*, 1975, pp. 86–98.

Stiglitz, Joseph E., "Pareto efficient taxation and expenditures: Pre- and re-distribution," *Journal of Public Economics*, 2018, *162*, 101–119. In Honor of Sir Tony Atkinson (1944-2017).

The Federal Reserve, "Distribution of Household Wealth in the U.S. since 1989," 2022.

Timperley, Jocelyn, "Should we give up flying for the sake of the climate?," 2020.

Toikka, Juuso, "Ironing without control," *Journal of Economic Theory*, 2011, *146* (6), 2510–2526.

Tullock, Gordon, "Excess benefit," *Water Resources Research*, 1967, *3* (2), 643–644.

Appendix A. Proofs

We now provide the omitted proofs. For the reader's convenience, we restate the results inline before providing the proofs.

Proposition 1 (Pareto Efficiency)**.** *The following are equivalent:*

(i) (x, t) *is the Pigouvian tax with equal redistribution.*

(ii) (x, t) *is incentive compatible and Pareto efficient.*

(iii) (x, t) *is optimal for a utilitarian planner with Pareto weights* $1/\eta$*, i.e., maximizes* (4) *subject to* (B) *and* (IC)*.*

PROOF. ((iii) $\implies$ (i)) Associating the weight $1/\eta$ on an agent with type (θ, c, η) is the same as considering an agent with transformed type $(\frac{\theta}{\eta}, \frac{c}{\eta}, 1)$: let G denote the associated distribution of transformed types in society.

The objective (4) of a planner who puts a welfare weight of $1/\eta$ on an agent of type (θ, η, c) can be rewritten as

$$\max_{x,t} \left(\int \frac{\theta}{\eta} v(x) dG \right) - \mathbb{E}\left[\frac{c}{\eta}\right] d(\bar{x}) - \int t(\theta, c, \eta) dG.$$

To maximize this objective while ignoring the incentive constraints (IC), it is clearly optimal to choose the sum of transfers as small as possible subject to (B) which yields $\int t(\theta, c, \eta) dG = \kappa \bar{x}$. The necessary first order conditions (ignoring the incentive constraints) therefore are that:

$$x(\theta, c, \eta) \in \underset{z \in X}{\operatorname{argmax}} \frac{\theta}{\eta} v(z) - (\kappa + \mathbb{E}\left[c/\eta\right] d'(\bar{x})) z \,. \tag{21}$$

Note that the above allocation equals the allocation in a Pigouvian tax (see Definition 3). As (B) is satisfied the mechanism does not run a surplus. As the Pigouvian tax with equal redistribution is incentive compatible, it is an optimal mechanism in (B). As transfers are unique up to a constant in any incentive compatible mechanism implementing the same allocation, the Pigouvian tax is also the unique optimal mechanism.

((i) $\implies$ (ii)) Note that by (21), charging the linear tax $(\kappa + \mathbb{E}\left[c/\eta\right] d'(\bar{x}))$ and letting agents select their desired level of consumption achieves a Pareto efficient allocation. Further, equal redistribution implies incentive compatibility as the lump-sum transfer $-t(0)$ received by each agent does not depend on the agent's report.

((ii) $\implies$ (iii)) Any Pareto efficient allocation is also Pareto efficient for the transformed types $(\frac{\theta}{\eta}, \frac{c}{\eta}, 1)$, which are distributed according to G. A society where types are distributed according to G is one where the agents' preferences are all quasilinear in transfers and have equal value for money. Pareto efficiency in a quasilinear society with equal value for money is equivalent to utilitarian efficiency. Therefore, a Pareto efficient outcome must solve:

$$\max_{x,t} \int U\left(\frac{\theta}{\eta}, \frac{c}{\eta}, 1; x, t\right) dG\left(\frac{\theta}{\eta}, \frac{c}{\eta}, 1\right), \tag{Pareto}$$
$$\text{s.t. (B) and (IC).}$$

As this is the same problem as maximizing (4) subject to (B) and (IC) the result follows. ■

OBSERVATION 1. *If $\eta = 1$ a.s. then the Pigouvian tax (i) is Pareto efficient, (ii) allocative efficient, and (iii) utilitarian efficient.*

PROOF. This follows straightforwardly from Proposition 1 when $\eta = 1$ a.s. In particular, since $\eta = 1$ is a special case of the general model, Pigouvian taxes are Pareto efficient. The proof of the proposition shows that this is utilitarian efficient. Finally, observe that since $\eta = 1$ a.s., (21) implies the resulting allocation is allocative efficient. ■

COROLLARY 1 (aggregate-consumption-constrained Pareto Efficiency)**.** *The following are equivalent:*

(i) (x, t) *is a linear tax mechanism with equal redistribution.*
(ii) (x, t) *is incentive compatible and aggregate-consumption-constrained Pareto efficient.*
(iii) (x, t) *is optimal for a utilitarian planner with Pareto weight* $1/\eta$ *over all mechanisms that induce aggregate consumption* $\bar{x}$*, i.e. maximizes* (4) *subject to* (B) *and* (IC).

We omit the proof as Corollary 1 follows from the same argument as Proposition 1.

PROPOSITION 2 (Incentive Compatibility)**.** *A mechanism* (x, t) *is incentive compatible if and only if*

(i) (x, t) *a.s. depends only on* $\rho = \theta/\eta$;
(ii) the allocation x *is non-decreasing in* ρ*; and*
(iii) the transfer as a function of the ratio ρ *satisfies*

$$t(\rho) = \rho v(x(\rho)) - \int_0^\rho v(x(r))dr + t(0)\,. \tag{5}$$

PROOF. We first observe that $\bar{x} = \int x(\theta, c, \eta)dF$ does not change if the type of any individual agent changes. Thus no agent's report affects the total consumption in society. Incentive compatibility is equivalent to the condition that for each pair of types (θ, c, η) and (θ', c', η')

$$v(x(\theta, c, \eta))\frac{\theta}{\eta} - t(\theta, c, \eta) \geq v(x(\theta', c', \eta'))\frac{\theta}{\eta} - t(\theta', c', \eta')\,.$$

That x must be non-decreasing follows immediately from the fact that v is strictly increasing, and the usual standard arguments; and is omitted.

To simplify notation we define $V(\rho, c, \eta) = \rho v(x(\rho\eta, c, \eta)) - t(\rho\eta, c, \eta)$. For a pair of types (θ, c, η) and (θ', c', η') with $\rho = \theta/\eta = \theta'/\eta'$ incentive compatibility implies that

$$V(\rho, c, \eta) = v(x(\theta, c, \eta))\frac{\theta}{\eta} - t(\theta, c, \eta) = v(x(\theta', c', \eta'))\frac{\theta}{\eta} - t(\theta', c', \eta') = V(\rho, c', \eta')\,.$$

By the envelope theorem (Theorem 2 in Milgrom and Segal, 2002) we have that V is absolutely continuous and for any c, η, c', η' and for Lesbegue almost every ρ[39]

$$v(x(\rho\eta, c, \eta)) = V_\rho(\rho, c, \eta) = V_\rho(\rho, c', \eta') = v(x(\rho\eta', c', \eta')).$$

Since v is strictly increasing, it follows that x depends only on ρ (up to a set of Lebesgue measure zero). Incentive compatibility immediately implies that the transfer must almost surely depend only on ρ: if two reports would lead to the same allocation but a different transfer, it would be optimal to deviate by misreporting η to minimize the payment. We can hence write $x(\rho), t(\rho)$ for the allocation and transfer as a function only of ρ. As $V_\rho(\rho, c, 1) = v(x(\rho))$ and $V(0, c, 1) = -t(0)$ we have that

$$V(\rho, c, 1) = v(x(\rho))\rho - t(\rho) = \int_0^\rho v(x(z))dz - t(0)$$

which implies (5). ■

OBSERVATION 2. *There exists an optimal mechanism for the planner such that the budget feasibility constraint* (B) *binds, i.e.,*

$$\int t(\rho)dH = \kappa \int x(\rho)dH\,.$$

Further, (B) *binds for any optimal mechanism whenever* $\alpha < 1$.

PROOF. To see this, recall that the planner's objective function is

$$\alpha \int (\theta v(x(\theta, c, \eta)) - \kappa x(\theta, c, \eta) - cd(\bar{x}))\, dF + (1 - \alpha) \int U(\theta, c, \eta; x, t) dF\,.$$

Observe that if $\alpha = 1$, then transfers do not enter into the planner's objective function, and the budget feasibility constraint can be taken to be binding without loss: any money collected from the agents can be rebated to them without affecting incentives (since there is a mass of agents, any single agent's transfer does not affect the rebate).

Now consider $\alpha < 1$. In this case, recall that $U(\theta, c, \eta; x, t)$ is defined as:

$$U(\theta, c, \eta; x, t) = \theta\, v(x(\theta, c, \eta)) - cd(\bar{x}) - \eta\, t(\theta, c, \eta)\,.$$

Further, in light of Proposition 2, we have that (x, t) can be written as functions of $\rho = \theta/\eta$, and further that

$$U(\theta, c, \eta; x, t) = -cd(\bar{x}) + \eta \int_0^{\theta/\eta} v(x(g))dg - \eta t(0).$$

[39] We denote derivatives by subindices.

Therefore the objective function of the planner is decreasing in $t(0)$. The budget constraint (B) requires that

$$\int t(\rho)dH \geq \int \kappa x(\rho)dH\,,$$

which in light of Proposition 2 can be rewritten as:

$$t(0) \geq \int \left(-zv(x(z)) + \int_0^z v(x(z))dz + \kappa x(z)\right) dH.$$

Fixing x, by observation, the planner is made better off, i.e. $\alpha V_A + (1-\alpha)V_U$ can be increased by making $t(0)$ as small as possible (such that (B) binds). The proposition follows. ■

PROPOSITION 3. *The welfare of the designer in a budget-balanced, incentive compatible mechanism (x,t) is given by*

$$\int (J(\rho)v(x(\rho)) - \kappa x(\rho))\, dH - \mathbb{E}\left[c\right] d(\bar{x}) - \kappa\bar{x}, \tag{6}$$

where the "virtual value" $J : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$J(z) = \alpha\, \mathbb{E}\left[\theta | \rho = z\right] + (1-\alpha)\left(z - (1 - \mathbb{E}\left[\eta \mid \rho \geq z\right]) \frac{1-H(z)}{h(z)}\right). \tag{7}$$

PROOF. The welfare of the planner is given by

$$\alpha \int (\theta v(x(\theta,c,\eta)) - \kappa x(\theta,c,\eta) - cd(\bar{x}))\, dF + (1-\alpha)\int U(\theta,c,\eta;x,t)dF\,.$$

Substituting in $U(\theta,c,\eta;x,t)$,

$$= \int ((\theta v(x(\theta,c,\eta)) - cd(\bar{x})) - \alpha\kappa x(\theta,c,\eta) - (1-\alpha)\eta t(\theta,c,\eta))\, dF,$$

$$= \int \eta\left(\frac{\theta}{\eta}x(\theta,c,\eta) - (1-\alpha)t(\theta,c,\eta)\right) - \alpha\kappa x(\theta,c,\eta)dF - \mathbb{E}\left[c\right] d(\bar{x}).$$

By Proposition 2 (x,t) are a.s. functions of ρ and therefore,

$$= \int \eta\left(\rho v(x(\rho)) - (1-\alpha)t(\rho)\right) dF - \alpha\kappa\bar{x} - \mathbb{E}\left[c\right] d(\bar{x}),$$

$$= \int \mathbb{E}\left[\eta|\rho = z\right] (zv(x(z)) - (1-\alpha)t(z))\, dH\; - \alpha\kappa\bar{x} - \mathbb{E}\left[c\right] d(\bar{x}).$$

Define $m(z) = \mathbb{E}\left[\eta|\rho = z\right]$,

$$= \int (m(z)zv(x(z)) - (1-\alpha)m(z)t(z))\, dH\; - \alpha\kappa\bar{x} - \mathbb{E}\left[c\right] d(\bar{x}).$$

From Proposition 2, recall that we have that:

$$t(\rho) = t(0) + \rho v(x(\rho)) - \int_0^{\rho} v(x(g))dg.$$

Substituting in, we have the planner's value is given by

$$\begin{aligned}
&\int (m(z)zv(x(z)) - (1-\alpha)m(z)t(z))\, dH - \alpha\kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}), \\
&= \int \left(\alpha m(z)zv(x(z)) + (1-\alpha)m(z)\int_0^z v(x(g))dg\right) dH - (1-\alpha)\mathbb{E}[m(z)]\, t(0) \\
&\quad - \alpha\kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}), \\
&= \int \left(\alpha m(z)z + (1-\alpha)\frac{\int_z^{\infty} m(g)dH}{h(z)}\right) v(x(z))dH - (1-\alpha)\mathbb{E}[m(\rho)]t(0) \\
&\quad - \alpha\kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}), \\
&= \int \left(\alpha m(z)z + (1-\alpha)\mathbb{E}[\eta|\rho \geq z]\frac{1-H(z)}{h(z)}\right) v(x(z))dH - (1-\alpha)\mathbb{E}[\eta]t(0) \\
&\quad - \alpha\kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}).
\end{aligned}$$

From Observation 2, applying the usual integration by parts argument, we have that, in any optimal mechanism,

$$t(0) = -\int \left(\rho - \frac{1-H(\rho)}{h(\rho)}\right) v(x(\rho))dH + \kappa\bar{x}.$$

Substituting in the above, and also substituting in our normalization that $\mathbb{E}[\eta] = 1$, we have the planner's objective function can be written as:

$$\begin{aligned}
&\int \left(\alpha m(z)z + (1-\alpha)\mathbb{E}[\eta|\rho \geq z]\frac{1-H(z)}{h(z)}\right) v(x(z))dH \\
&\qquad + (1-\alpha)\int \left(z - \frac{1-H(z)}{h(z)}\right) v(x(z))dH - \kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}), \\
&= \int \left(\alpha m(z)z + (1-\alpha)\left(z - (1-\mathbb{E}[\eta|\rho \geq z])\frac{1-H(z)}{h(z)}\right)\right) v(x(z))dH(z) \\
&\qquad - \kappa\bar{x} - \mathbb{E}[c]\, d(\bar{x}).
\end{aligned}$$

Noting that $m(z)z = \mathbb{E}[\eta|\rho = z]\, z = \mathbb{E}[\eta z|\rho = z] = \mathbb{E}[\theta|\rho = z]$, we have the desired formula. ■

LEMMA 1. *The virtual value $J : [0, \overline{\theta}/\underline{\eta}] \to \mathbb{R}$ has the following properties:*

(i) J is continuous,

(ii) $J(0) = 0$,

(iii) $J(\overline{\theta}/\underline{\eta}) = \overline{\theta}(\alpha + (1-\alpha)1/\underline{\eta}) \geq \overline{\theta}$.

PROOF. The continuity of J follows as all of

$$h(z) = \int\int f(z\eta, c, \eta) dc d\eta,$$
$$\mathbb{E}\left[\theta|\rho = z\right] = \frac{\int\int \rho\eta f(\rho\eta, c, \eta) d\eta dc}{h(\rho)},$$
$$\mathbb{E}\left[\eta|\rho \geq z\right] = \frac{\int\int\int \eta \mathbf{1}_{\theta/\eta \geq z} f(\theta, c, \eta) d\theta d\eta dc}{1 - H(\rho)}$$

are continuous functions of ρ by the continuity of f.

By observation we have that $\mathbb{E}\left[\theta|\rho = 0\right] = 0$. Furthermore, $\mathbb{E}\left[\eta|\rho \geq 0\right] = \mathbb{E}\left[\eta\right]$ which by our normalization equals 1, Together this implies that $J(0) = 0$.

Next, we have that $\mathbb{E}\left[\theta\middle|\rho = \overline{\theta}/\underline{\eta}\right] = \overline{\theta}$ and $H(\overline{\theta}/\underline{\eta}) = 1$ which together imply that $J(\overline{\theta}/\underline{\eta}) = \overline{\theta}(\alpha + (1-\alpha)1/\underline{\eta})$. ■

THEOREM 1. *The optimal mechanism (x, t) satisfies:*

$$x(\rho) \in \underset{x \in X}{\operatorname{argmax}}\, \bar{J}(\rho) v(x) - (\mathbb{E}\left[c\right] d'(\bar{x}) + \kappa) x\,. \tag{10}$$

With transfer equal to

$$t(\rho) - t(0) = \rho v(x(\rho)) - \int_0^\rho v(x(r)) dr, \tag{11}$$

and the redistributive transfer $t(0)$ is given by

$$t(0) = -\int_0^{\overline{\theta}/\underline{\eta}} \left(\left[\rho - \frac{1 - H(\rho)}{h(\rho)}\right] v(x(\rho)) - \kappa x(\rho) \right) dH(\rho)\,.$$

PROOF. Formally, recall that our problem can be written as:

$$\max_{x(\cdot)} \int J(z) v(x(z)) dH(z) - \kappa \int x(z) dH(z) - \mathbb{E}\left[c\right] d\left(\int x(z) dH(z)\right), \tag{22}$$
$$\text{s.t. } x(\cdot) \text{ is non-decreasing.}$$

Let x^* be a solution to this problem, with $\bar{x}^*$ the associated aggregate consumption. By convexity of $d(\cdot)$ and concavity of v, x^* is also optimal in

$$\max_{x(\cdot)} \int J(z) v(x(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) x(z) dH(z), \tag{23}$$
$$\text{s.t. } x(\cdot) \text{ is non-decreasing.}$$

To see this suppose not, i.e., suppose there exists x' non-decreasing such that

$$\int J(z)v(x'(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) x'(z)dH(z) > \int J(z)v(x^*(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) x^*(z)dH(z),$$

Then note that by concavity of v for any $\alpha \in [0,1]$, $x'' = \alpha x' + (1-\alpha)x^*$,

$$\int J(z)v(x''(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) x''(z)dH(z) > \int J(z)v(x^*(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) x^*(z)dH(z).$$

This implies that, evaluated at $\alpha = 0$,

$$\frac{d}{d\alpha}\left(\int J(z)v(\alpha x'(z) + (1-\alpha)x^*(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) (\alpha x'(z) + (1-\alpha)x^*(z))dH(z)\right) > 0,$$

$$\implies \int J(z)v'(x^*(z))(x'(z) - x^*(z)) - \left(\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*)\right) (x'(z) - x^*(z))dH(z) > 0.$$

However this contradicts the optimality of x^* in the original problem.

So now consider the problem (23). If $J(\cdot)$ is not non-decreasing, let $\bar{J}(\cdot)$ be the ironed J defined in text. For each ρ, define the point-wise maximization problem

$$\max_x \bar{J}(\rho)v(x) - (\kappa + \mathbb{E}\left[c\right] d'(\bar{x}^*))x.$$

By Corollary 3.9 of Toikka (2011), the solution to this pointwise problem must also solve (23), and therefore by the argument above also the solution to (22). Then, (11) follows from Proposition 2 and the characterization of redistributive transfer $t(0)$ follows from Observation 2. ■

COROLLARY 2 (Linear Environments). *Suppose $X = [0,1]$ and $v(x) = d(x) = x$.*

(i) It is optimal to prohibit consumption of the good iff $\bar{J}(\rho) \leq \mathbb{E}\left[c\right] + \kappa$ a.s..

(ii) Otherwise, charging a constant price per unit of consumption with full redistribution is optimal, and the optimal price ρ^* solves $\bar{J}(\rho^*) = \mathbb{E}\left[c\right] + \kappa$.

PROOF. By Proposition 2, and Observation 2 there exists an optimal mechanism that runs no surplus, where the allocation depends only on the ratio $\rho = \theta/\eta$. Hence, in a linear environment, any mechanism (x,t), such that the allocation $x : [0, \overline{\theta}/\underline{\eta}] \to [0,1]$ solves

$$\max_x \int x(\rho)(J(\rho) - (\kappa + c))dH \tag{24}$$

$$\text{s.t. } x \text{ non-decreasing }, \tag{25}$$

and the transfer t solves (5) constitutes an optimal mechanism. All extreme points of the set of non-decreasing functions from $[0, \overline{\theta}/\underline{\eta}]$ to $[0,1]$ are of the form $x(\rho) = \mathbf{1}_{\rho \geq \rho^*}$ and thus correspond to a posted-price mechanism. As (24) is linear in x, there exists an extreme point that is a maximizer, by Bauer's Maximum principle.

We can thus rewrite the designer's optimization problem as an optimization problem over reserve prices ρ^*

$$\max_{\rho^*} \int_{\rho^*}^{\overline{\theta}/\underline{\eta}} (J(\rho) - (\kappa + c)) dH \,.$$

The planner's objective is differentiable with respect to ρ^* with derivative

$$-(J(\rho^*) - (\kappa + c))h(\rho) \,.$$

If $J(\rho) \leq c$ for all $\rho \in [0, \overline{\theta}/\underline{\eta}]$ the derivative is always positive, and any reserve price above $\overline{\theta}/\underline{\eta}$ that excludes almost all types from consumption of the good is optimal. Otherwise, as J is continuous by Lemma 1, the derivative is continuous and hence the first-order condition implies that $J(\rho^*) = c + \kappa$. ∎

PROPOSITION 4. *Suppose that $X = \mathbb{R}_+$ and v is strictly concave. The optimal allocation can be implemented by a tax that depends on consumption given by*[40]

$$t(x) - t(0) = (\kappa + \mathbb{E}\,[c]\, d'(\bar{x}^*)) \int_0^x \frac{(x^*)^{-1}(q)}{\bar{J}((x^*)^{-1}(q))}\, dq, \textit{ where,} \tag{13}$$

$$x^*(\rho) = v'^{-1}\left(\frac{\kappa + \mathbb{E}\,[c]\, d'(\bar{x}^*)}{\bar{J}(\rho)} \right), \tag{14}$$

$$\bar{x}^* = \int x^*(\rho) dH,$$

PROOF. Under the above nonlinear consumption tax the agents' optimal consumption $\tilde{x}$ satisfies

$$\rho v'(\tilde{x}(\rho)) = (\kappa + \mathbb{E}\,[c]\, d'(\bar{x})) \frac{x^{*-1}(\tilde{x}(\rho))}{\bar{J}(x^{*-1}(\tilde{x}(\rho)))} \,.$$

We note that this equation is uniquely solved by $\tilde{x}(\rho) = x^*(\rho)$ for $\bar{J}(\rho) \geq \bar{x}^*$.

Furthermore, by observation, we have that $t'(0) = (\kappa + \mathbb{E}\,[c]\, d'(\bar{x}^*)) \frac{r(0)}{\bar{J}(r(0))}$. Using the fact that $v'(0) = 1$ we have that $\bar{J}(r(0)) = (\kappa + \mathbb{E}\,[c]\, d'(\bar{x}^*))$. Therefore $\tilde{x}(\rho) = 0$ whenever $\rho \leq r(0)$. ∎

COROLLARY 3. *Suppose that $X = \mathbb{R}_+$ and v is strictly concave and differentiable. If $\bar{J}(z)/z$ is increasing (decreasing), then the marginal tax on consumption is decreasing (respectively, increasing). A linear tax (constant marginal tax) on consumption is optimal if and only if J is linear.*

[40] Here again, $(x^*)^{-1}$ denotes the pseudo inverse defined as $(x^*)^{-1}(q) = \sup\{\rho \colon x^*(\rho) \leq q\}$.

PROOF. To see this recall that from (13), we have that

$$t(x) - t(0) = (\kappa + \mathbb{E}[c]\, d'(\bar{x}^*)) \int_0^x \frac{(x^*)^{-1}(q)}{\bar{J}((x^*)^{-1}(q))} dq,$$
$$\implies t'(x) = \frac{(x^*)^{-1}(x)}{\bar{J}((x^*)^{-1}(x))}$$

Note that $(x^*)^{-1}(\cdot)$ is an increasing function. Therefore we have that the marginal tax is increasing (respectively, decreasing) whenever $z/\bar{J}(z)$ is increasing (respectively, decreasing), as desired.

Finally, to see that J being linear is necessary for a linear tax/ constant marginal tax, note that it must be that $z/\bar{J}(z)$ is constant, and therefore $\bar{J}(\rho)$ is a linear function of ρ. Note that in this case $J(\rho)$ must be a linear function of ρ since $\bar{J} = J$ in any neighborhood where $\bar{J}$ is not locally constant.

■

PROPOSITION 5. *Consider a utilitarian planner. Assume $d(x) = x$ and J non-decreasing when J is positive (so that no ironing is necessary).*

(i) If $\lim_{\rho\to\infty}(1 - H(\rho))/h(\rho) = 0$ and $Cov(c, 1/\eta) \geq 0$, then there exists ρ' such that the consumption of agents with willingness to pay $\rho \geq \rho'$ is strictly higher in the optimal mechanism than under Pigouvian taxes.

(ii) Let $\mathbb{E}[\eta|\rho = 0] > 2$. For $\kappa, \mathbb{E}[c/\eta]$ low enough, there exists ρ'' that the consumption of agents with willingness to pay $\rho \leq \rho''$ is weakly lower in the optimal mechanism than under Pigouvian taxes and strictly lower for some ρ.

PROOF. To see (i), observe that substituting the assumption that the inverse rate vanishes into (7), we have that the virtual value converges to weakly larger than the WTP, i.e., $\lim_{\rho\to\infty} J(\rho)/\rho \geq 1$. Since $\mathrm{Cov}(c, 1/\eta) \geq 0$ we have that $\mathbb{E}[c/\theta] > \mathbb{E}[c]$. Therefore for ρ large we have that $\frac{\mathbb{E}[c]+\kappa}{J(\rho)} < \frac{\mathbb{E}[c/\theta]+\kappa}{\rho}$.

To see (ii) note that

$$J(z)h(z) = h(z)z - (1 - H(z)) + \mathbb{E}\left[\eta \mathbb{1}_{\{\rho \geq z\}}\right]$$

This yields that

$$(J(z)h(z))'|_{z=0} = 2h(0) + \frac{\partial}{\partial z}\mathbb{E}\left[\eta \mathbb{1}_{\{\rho \geq z\}}\right] = 2h(0) + \frac{\partial}{\partial z}\int_z^\infty \mathbb{E}[\eta|\rho = y]\, h(y)dy$$
$$= 2h(0) - \mathbb{E}[\eta|\rho = 0]\, h(0) = h(0)\left(2 - \mathbb{E}[\eta|\rho = 0]\right).$$

As the above term is negative and J is continuous by Lemma 1, J is strictly negative for an interval around 0. As J is strictly negative in an interval around 0 the optimal mechanism allocates 0 consumption to agents in this interval. Under Pigouvian taxation, consumption is 0 for any type with willingness to pay $\rho \leq \kappa + \mathbb{E}[c/\eta]$, and positive for any type above. For $\mathbb{E}[c\eta], \kappa$ small enough, the interval where J is negative has a nonempty intersection with the set of agents who consume a strictly positive amount under Pigouvian taxation. The result follows. ■

PROPOSITION 6. *Consider a planner concerned with narrow allocative efficiency. Assume that $d(x) = x$, J is non-decreasing (so that no ironing is necessary), and $\mathbb{E}[\eta|\rho = z]$ is decreasing in z. Then, in the optimal mechanism for an allocative efficient planner:*

(i) If $\lim_{z\to\infty} \mathbb{E}[\eta|\rho = z] \to 0$, there exists $\bar{\rho}$ such that the consumption of agents with WTP $\rho \geq \bar{\rho}$ is strictly lower than under Pigouvian taxes.

(ii) If $Cov(c, 1/\eta) \geq 0$ and $\kappa, \mathbb{E}[c]$ are low enough, there exists $\underline{\rho}$ such that the consumption of agents with WTP to pay $\rho < \underline{\rho}$ is weakly higher than under Pigouvian taxes.

PROOF. To begin, note that for an allocative efficient planner, the virtual value (7) reduces to $J(z) = \mathbb{E}[\theta|\rho = z] = z\, \mathbb{E}[\eta|\rho = z]$.

To see (i), note that for for z large enough, we have that $J(z)/(\kappa + \mathbb{E}[c]) = (z\mathbb{E}[\eta|\rho = z])/(\kappa + \mathbb{E}[c])$ is smaller than $z/(\kappa + \mathbb{E}[c/\eta])$ since $\lim_{z\to\infty} \mathbb{E}[\eta|\rho = z] = 0$. The result follows.

Conversely, for (ii), note that for z small, we must have $\mathbb{E}[\eta|\rho = z] \geq 1$ since $\mathbb{E}[\eta|\rho = z]$ is decreasing in z and $\mathbb{E}[\eta] = 1$, and therefore $J(z) > z$. Further, since $\text{Cov}(c, 1/\eta) \geq 0$, $\mathbb{E}[c/\eta] > \mathbb{E}[c]$. Therefore, for z small enough, an allocative efficient planner would allocate the agent weakly more than under Pigouvian taxation. If κ and $\mathbb{E}[c]$ are small enough, the allocation would be strictly greater for WTP $z > \kappa + \mathbb{E}[c]$. ■

PROPOSITION 7. *Suppose the agent's value for the good and value for money are both distributed according to the Rayleigh distribution as described above. Then, we have that*

$$J(z) = \alpha \frac{3z\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}} + (1-\alpha)\left(z - \left(1 - \frac{\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}}\right)\frac{z^2\sigma_\eta^2 + \sigma_\theta^2}{2z\sigma_\eta^2}\right). \tag{15}$$

No ironing is needed in the optimal mechanism, i.e., $\overline{J}(z) = J(z)$*. Further, if* v *is strictly concave, the tax on consumption that implements the optimal mechanism is such that marginal taxes are increasing in consumption, for either objective of the planer.*

PROOF. Let's denote the PDF and CDF of θ by f and F; and similarly the PDF and CDF of η by g and G. We're interested in the distribution of $\rho = \theta/\eta$, which we'll denote by h, H. To calculate this, observe that:

$$\begin{aligned}
1 - H(r) &= \mathbb{P}\left[\frac{\theta}{\eta} \geq r\right] = \mathbb{P}\left[\theta \geq r\eta\right] = \mathbb{E}_\eta[1 - F(r\eta)] = \mathbb{E}_\eta\left[e^{-\frac{r^2\eta^2}{2\sigma_\theta^2}}\right], \\
&= \int_0^\infty \frac{\eta}{\sigma_\eta^2} e^{-\frac{r^2\eta^2}{2\sigma_\theta^2}} e^{-\frac{\eta^2}{2\sigma_\eta^2}} d\eta = \int_0^\infty \frac{\eta}{\sigma_\eta^2} e^{-\frac{(r^2\sigma_\eta^2 + \sigma_\theta^2)\eta^2}{2\sigma_\theta^2\sigma_\eta^2}} d\eta = \frac{\sigma_\theta^2}{r^2\sigma_\eta^2 + \sigma_\theta^2}, \\
h(r) &= \frac{2r\sigma_\theta^2\sigma_\eta^2}{(r^2\sigma_\eta^2 + \sigma_\theta^2)^2}.
\end{aligned}$$

A quick calculation verifies that ρ has a finite expectation when distributed according to h, H. Therefore

$$\rho - \frac{1 - H(\rho)}{h(\rho)} = \rho - \frac{\rho^2\sigma_\eta^2 + \sigma_\theta^2}{2\rho\sigma_\eta^2} = \frac{\rho}{2} - \frac{\sigma_\theta^2}{2\rho\sigma_\eta^2},$$

which is increasing in ρ by observation.

Recall next from (7) that we have that:

$$J(z) = \alpha\, \mathbb{E}\left[\theta | \rho = z\right] + (1-\alpha)\left(z - (1 - \mathbb{E}\left[\eta | \rho \geq z\right])\frac{1 - H(z)}{h(z)}\right).$$

So the remaining terms we need to compute are $\mathbb{E}\left[\theta|\rho = z\right]$ and $\mathbb{E}\left[\eta|\rho \geq z\right]$. Let's try to do these in turn. We will stick to the notation that f, F is the density/CDF corresponding to θ, g, G corresponding to η and h, H corresponding to ρ. Let l, L correspond to the joint density, CDF of (θ, ρ). Observe that:

$$L(\theta, \rho) = \int_0^\theta f(t)\left(1 - G\left(\frac{t}{\rho}\right)\right) dt, \implies l(\theta, \rho) = \frac{\theta}{\rho^2} f(\theta) g\left(\frac{\theta}{\rho}\right).$$

Therefore, we can write:

$$\mathbb{E}\left[\theta|\rho=z\right]=\frac{\int_0^\infty \theta l(\theta,z)d\theta}{h(z)}=\frac{1}{h(z)}\int_0^\infty \frac{\theta^2}{z^2}f(\theta)g\left(\frac{\theta}{z}\right)d\theta,$$
$$=\frac{1}{h(z)}\int_0^\infty \frac{\theta^4}{z^3\sigma_\theta^2\sigma_\eta^2}\exp\left\{-\frac{\theta^2(z^2\sigma_\eta^2+\sigma_\theta^2)}{2z^2\sigma_\theta^2\sigma_\eta^2}\right\}d\theta.$$

Define $\sigma^2=\frac{z^2\sigma_\theta^2\sigma_\eta^2}{(z^2\sigma_\eta^2+\sigma_\theta^2)}$.

$$=\frac{1}{h(z)}\frac{1}{z(\sigma_\theta^2+z^2\sigma_\eta^2)}\int_0^\infty \theta^3\left(\frac{\theta}{\sigma^2}\right)\exp\left\{-\frac{\theta^2}{2\sigma^2}\right\}d\theta,$$

Letting $m(\theta)$ be the density of a Rayleigh distribution with parameter σ^2, we have:

$$=\frac{1}{h(\rho)}\frac{1}{\rho(\sigma_\theta^2+\rho^2\sigma_\eta^2)}\int_0^\infty \theta^3 m(\theta)d\theta=\frac{1}{\sigma^2}\int_0^\infty \theta^3 m(\theta)d\theta=\frac{1}{\sigma^2}\sigma^3\frac{3\sqrt{\pi}}{\sqrt{2}},$$
$$=\frac{3\sigma\sqrt{\pi}}{\sqrt{2}}.$$

Substituting in σ from above we have

$$\mathbb{E}\left[\theta|\rho=z\right]=\frac{3\sqrt{\pi}z\sigma_\theta\sigma_\eta}{\sqrt{2(z^2\sigma_\eta^2+\sigma_\theta^2)}}.$$

Similarly,

$$\mathbb{E}\left[\eta|\rho\geq z\right]=\frac{\int_0^\infty \eta g(\eta)(1-F(z\eta))d\eta}{1-H(z)}=\frac{1}{1-H(z)}\int_0^\infty \frac{\eta^2}{\sigma_\eta^2}e^{-\frac{(z^2\sigma_\eta^2+\sigma_\theta^2)\eta^2}{2\sigma_\theta^2\sigma_\eta^2}}d\eta.$$

Defining $\sigma^2=\frac{\sigma_\eta^2\sigma_\theta^2}{z^2\sigma_\eta^2+\sigma_\theta^2}$, we have

$$=\int_0^\infty \frac{\eta^2}{\sigma^2}\exp\{-\frac{\eta^2}{2\sigma^2}\}d\eta.$$

By inspection this is simply the expectation of a Rayleigh distributed random variable with parameter σ^2, i.e.,

$$\mathbb{E}\left[\eta|\rho\geq z\right]=\frac{\sqrt{\pi}\sigma_\eta\sigma_\theta}{\sqrt{2(z^2\sigma_\eta^2+\sigma_\theta^2)}}.$$

So in summary, we have:

$$\frac{1-H(z)}{h(z)} = \frac{z}{2} + \frac{\sigma_\theta^2}{2z\sigma_\eta^2},$$
$$\mathbb{E}\left[\eta | \rho \geq z\right] = \frac{\sqrt{\pi}\sigma_\eta\sigma_\theta}{\sqrt{2(z^2\sigma_\eta^2 + \sigma_\theta^2)}},$$
$$\mathbb{E}\left[\theta | \rho = z\right] = \frac{3\sqrt{\pi}\rho\sigma_\theta\sigma_\eta}{\sqrt{2(z^2\sigma_\eta^2 + \sigma_\theta^2)}}.$$

Since $\sigma_\eta = \sqrt{2/\pi}$ by our normalization, we can calculate $J(z)$ is as desired.

To see that $\bar{J}(z) = J(z)$, note that for $\alpha = 1$,

$$J(z) = \frac{3z\sigma_\theta}{\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)}} = \frac{3\sigma_\theta}{\sqrt{(\sigma_\eta^2 + (\sigma_\theta^2/z^2))}},$$

which is increasing by observation. Next for $\alpha = 0$

$$\begin{aligned}
J(z) &= z - \left(1 - \frac{\sigma_\theta}{\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)}}\right)\frac{z^2\sigma_\eta^2 + \sigma_\theta^2}{2z\sigma_\eta^2}, \\
&= \frac{2z^2\sigma_\eta^2 - (\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)} - \sigma_\theta)\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)}}{2z\sigma_\eta^2}, \\
&= \frac{2z\sigma_\eta^2 - (\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)} - \sigma_\theta)\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}}{2\sigma_\eta^2}. \\
\implies J'(z) &= \frac{1}{2\sigma_\eta^2}\left(2\sigma_\eta^2 - 2\sigma_\eta^2 + \frac{\sigma_\theta^2(\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)} - \sigma_\theta)\sigma_\theta^2}{z^3\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}}\right), \\
&> 0.
\end{aligned}$$

Therefore since $J(\cdot)$ is increasing for $\alpha = 0$ and $\alpha = 1$, it is also increasing for all $\alpha \in [0,1]$..

Finally, we are left to show that marginal taxes are increasing. Observe that for $\alpha = 1$,

$$\frac{J(z)}{z} = \frac{3\sigma_\theta}{\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)}},$$

which is decreasing in z. Next, consider $\alpha = 0$,

$$\frac{J(z)}{z} = \frac{2z^2\sigma_\eta^2 - (\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)} - \sigma_\theta)\sqrt{(z^2\sigma_\eta^2 + \sigma_\theta^2)}}{2z^2\sigma_\eta^2},$$

$$= \frac{2\sigma_\eta^2 - \left(\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)} - \sigma_\theta/z\right)\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}}{2\sigma_\eta^2}.$$

Therefore $J(z)/z$ is decreasing iff $\left(\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)} - \sigma_\theta/z\right)\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}$ is increasing in z. For brevity, define $w = \sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}$. Note that $dw/dz = -\frac{\sigma_\theta^2}{wz^3}$. Now, observe that:

$$\begin{aligned}
&\frac{d}{dz}\left(\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)} - \sigma_\theta/z\right)\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)} \\
=&\frac{d}{dz}\left(w - \sigma_\theta/z\right)w \\
=&\left(\frac{dw}{dz} + \sigma_\theta/z^2\right)w + \left(w - \sigma_\theta/z\right)\frac{dw}{dz} \\
=&\left(-\frac{\sigma_\theta^2}{wz^3} + \sigma_\theta/z^2\right)w - \frac{\sigma_\theta^2}{wz^3}\left(w - \sigma_\theta/z\right) \\
=&\left(w - \sigma_\theta/z\right)\frac{\sigma_\theta}{z^2} - \frac{\sigma_\theta^2}{wz^3}\left(w - \sigma_\theta/z\right) \\
=&\frac{\sigma_\theta}{wz^2}\left(w - \sigma_\theta/z\right)^2 \\
>&0,
\end{aligned}$$

as desired. Therefore, $J(z)/z$ is decreasing in z for all $\alpha \in [0, 1]$ and therefore marginal taxes are increasing by Corollary 3. ■

COROLLARY 4. *Suppose (θ, η) are distributed according to the Rayleigh distribution.*

(i) For a planner with the allocative objective ($\alpha = 1$), the optimal mechanism features a consumption cap, i.e., a maximum amount that is allowed to be consumed by any agent. If $\mathbb{E}[c]$ is small enough, some agent types consume more *than they would under Laissez-faire.*

(ii) For a planner with the utilitarian objective ($\alpha = 0$), the optimal mechanism has no consumption cap.

PROOF. We prove these in turn. Observe that for $\alpha = 1$, $J(z) = 3z\sigma_\theta/\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)} = 3\sigma_\theta/\sqrt{(\sigma_\eta^2 + \sigma_\theta^2/z^2)}$. It is then straightforward to verify that $J(\cdot)$ is strictly increasing. Observe that as $z \to \infty$, $J(z) \to 3\sigma_\theta/\sigma_\eta$. Then by Theorem 1, we have that the optimal allocation in this case solves $x(z) = v'^{-1}((\mathbb{E}[c]\, d'\bar{x} + \kappa)/J(z)$. Since $J(\cdot)$ is bounded from above, v'^{-1} is bounded from below and therefore since v is concave, $x(z)$ must be bounded from above, i.e. the optimal solution involves a consumption cap.

Next observe that for z small, $J(z) \approx 3z$. $J(0) = 0$ and $J(z) < z$ for z large, we have that $J(z) > z$ for $z \in (0, \bar{z})$. Depending on the values of c and κ the optimal allocation for (some of) these types will exceed their consumption under Laissez-Faire or Pigouvian taxes.

Next, for $\alpha = 0$, note that

$$J(z) = z - \left(1 - \frac{\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}}\right) \frac{z^2\sigma_\eta^2 + \sigma_\theta^2}{2z\sigma_\eta^2}.$$

By observation $J(z) < z$ since $\left(1 - \frac{\sigma_\theta}{\sqrt{(\sigma_\eta^2 z^2 + \sigma_\theta^2)}}\right) \geq 0$. Next, as $z \to \infty$, $J(z) - z/2 \to 0$ so $J(\cdot)$ is unbounded in this case. ■

REMARK 3. *If value for the good and value for money are jointly lognormally distributed as above, then they are jointly lognormally distributed with* $\rho = \frac{\theta}{\eta}$ *and*

$$\begin{pmatrix} \log\theta \\ \log\eta \\ \log\rho \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_\theta \\ -\sigma_\eta^2/2 \\ \mu_\rho \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \xi\sigma_\theta\sigma_\eta & \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta \\ \xi\sigma_\theta\sigma_\eta & \sigma_\eta^2 & \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2 \\ \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta & \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2 & \sigma_\rho^2 \end{pmatrix}\right),$$

where $\mu_\rho = \mu_\theta + \sigma_\eta^2/2$ *and* $\sigma_\rho^2 = \sigma_\theta^2 + \sigma_\eta^2 - 2\xi\sigma_\theta\sigma_\eta$.

PROOF. We note first ρ is lognormally distributed $\log\rho \sim \mathcal{N}(\mu_\rho, \sigma_\rho^2)$. To see this note firstly that $\log\rho = \log\theta - \log\eta$, and therefore $\log\rho$ is also normally distributed with $\mu_\theta + \sigma_\eta^2/2$. To verify the variance, observe that

$$\begin{aligned}
&\mathbb{E}\left[(\log\rho - \mu_\theta - \sigma_\eta^2/2)^2\right] = \mathbb{E}\left[([\log\theta - \mu_\theta] - [\log\eta + \sigma_\eta^2/2])^2\right], \\
&= \mathbb{E}\left[(\log\theta - \mu_\theta)^2\right] - 2\mathbb{E}\left[(\log\theta - \mu_\theta)(\log\eta + \sigma_\eta^2/2)\right] + \mathbb{E}\left[(\log\eta + \sigma_\eta^2/2)^2\right], \\
&= \sigma_\theta^2 + \sigma_\eta^2 - 2\kappa\sigma_\theta\sigma_\eta.
\end{aligned}$$

It is then straightforward to observe that all three are jointly lognormally distributed since the $\log\rho$ is linear combination of $\log\theta$ and $\log\eta$ which are already jointly normally distributed (and therefore any linear combination of the 3 is normally distributed)

The covariance between $\log\theta$ and $\log\rho$ equals $\sigma_\theta^2 - \kappa\sigma_\theta\sigma_\eta$. To see this note that

$$\begin{aligned}
&\text{Cov}\,(\log\theta, \log\rho), \\
&= \text{Cov}(\log\theta, \log\theta - \log\eta), && \text{(by definition of } \rho\text{)} \\
&= \text{Cov}(\log\theta, \log\theta) - \text{Cov}(\log\theta, \log\eta), && \text{(by definition of covariance)} \\
&= \sigma_\theta^2 - \kappa\sigma_\theta\sigma_\eta.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathrm{Cov}(\log\eta, \log\rho), \\
&= \mathrm{Cov}(\log\eta, \log\theta - \log\eta), && \text{(by definition of } \rho) \\
&= \mathrm{Cov}(\log\eta, \log\theta) - \mathrm{Cov}(\log\eta, \log\eta), && \text{(by definition of covariance)} \\
&= \kappa\sigma_\theta\sigma_\eta - \sigma_\eta^2.
\end{aligned}$$

Taken together, we have the statement of the remark as desired. ■

PROPOSITION 8. *Suppose the agent's value for the good and value for money are jointly lognormally distributed as in* (16)*. Then, the virtual value J is given by*

$$J(z) = \alpha\,\mathbb{E}[\theta]\,\frac{z^\beta}{\mathbb{E}\left[\rho^\beta\right]} + (1-\alpha)\left(z - \frac{H\left(ze^{(1-\beta)\sigma_\rho^2}\right) - H(z)}{h(z)}\right). \tag{18}$$

Here H, h are the CDF, PDF of ρ, and $\beta = Cov(\theta,\rho)/\sigma_\rho^2$. The "ironed virtual value" under the allocative objective $\alpha = 1$ and the utilitarian objective $\alpha = 0$ equals

$$\bar{J}(z) = \begin{cases} \mathbf{1}_{\beta\geq 0}J(z) + \mathbf{1}_{\beta<0}\mathbb{E}[\theta] & \text{for } \alpha = 1 \\ \max\{0, J(z)\} & \text{for } \alpha = 0\,. \end{cases}$$

PROOF. Let $\Phi_{\mu,\sigma}$ be the CDF and $\varphi_{\mu,\sigma}$ be the PDF of a Normal distribution with mean μ and standard deviation σ, and Φ, φ denote the PDF and CDF respectively of the standard Normal distribution, $\Phi = \Phi_{0,1}, \varphi = \varphi_{0,1}$.

Since θ and ρ are jointly lognormally distributed, i.e.,

$$\begin{pmatrix}\log\theta \\ \log\rho\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}\mu_\theta \\ \mu_\rho\end{pmatrix}, \begin{pmatrix}\sigma_\theta^2 & \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta \\ \sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta & \sigma_\rho^2\end{pmatrix}\right),$$

We therefore have that, conditional on ρ, θ is log normally distributed, i.e.

$$\log\theta|\rho = z \sim \mathcal{N}\left(\mu_\theta + \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2}(\log z - \mu_\rho), \left(1 - \frac{(\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta)^2}{\sigma_\theta^2\sigma_\rho^2}\right)\sigma_\theta^2\right).$$

Consequently,

$$\begin{aligned}
\mathbb{E}[\theta|\rho = z] &= \exp\left(\mu_\theta + \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2}(\log z - \mu_\rho) + \frac{1}{2}\left(1 - \frac{(\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta)^2}{\sigma_\theta^2\sigma_\rho^2}\right)\sigma_\theta^2\right), \\
&= \gamma_\theta z^{\beta_\theta}, \\
\text{where } \gamma_\theta &= \exp\left(\mu_\theta - \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2}\mu_\rho + \frac{1}{2}\left(1 - \frac{(\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta)^2}{\sigma_\theta^2\sigma_\rho^2}\right)\sigma_\theta^2\right),
\end{aligned}$$

$$\beta_\theta = \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2}.$$

Similarly, since η and ρ are jointly lognormally distributed, i.e.,

$$\begin{pmatrix}\log\eta\\ \log\rho\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}-\frac{1}{2}\sigma_\eta^2\\ \mu_\rho\end{pmatrix}, \begin{pmatrix}\sigma_\eta^2 & \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2\\ \xi\sigma_\theta\sigma_\eta - \sigma_\eta^2 & \sigma_\rho^2\end{pmatrix}\right).$$

Consequently, by analogous arguments to above, we have that, η is lognormally distributed conditional on ρ and therefore we have

$$\mathbb{E}\left[\eta|\rho = z\right] = \gamma_\eta z^{\beta_\eta},$$

$$\text{where } \gamma_\eta = \exp\left(-\sigma_\eta^2/2 - \frac{\xi\sigma_\theta\sigma_\eta - \sigma_\eta^2}{\sigma_\rho^2}\mu_\rho + \frac{1}{2}\left(1 - \frac{(\xi\sigma_\theta\sigma_\eta - \sigma_\eta^2)^2}{\sigma_\eta^2\sigma_\rho^2}\right)\sigma_\eta^2\right),$$

$$\beta_\eta = \frac{\xi\sigma_\theta\sigma_\eta - \sigma_\eta^2}{\sigma_\rho^2}.$$

Finally note that

$$\begin{aligned}\mathbb{E}\left[\eta|\rho \geq z\right] =& \mathbb{E}\left[\mathbb{E}\left[\eta|\rho = \hat{z}\right]|\hat{z} \geq z\right] = \mathbb{E}\left[\gamma_\eta \hat{z}^{\beta_\eta}|\hat{z} \geq z\right] = \gamma_\eta\mathbb{E}\left[\hat{z}^{\beta_\eta}|\hat{z} \geq z\right],\\ =& \gamma_\eta\mathbb{E}\left[\rho^{\beta_\eta}|\rho^{\beta_\eta} \geq z^{\beta_\eta}\right],\end{aligned}$$

Note that since ρ is distributed lognormally, i.e. $\log\rho \sim \mathcal{N}(\mu_\rho, \sigma_\rho^2)$, have that ρ^{β_η} is also distributed lognormally with parameters $(\beta_\eta\mu_\rho, \beta_\eta^2\sigma_\rho^2)$. Next, recall that for a lognormally distributed random variable τ s.t. $\log\tau \sim \mathcal{N}(\mu, \sigma^2)$, we have that

$$\mathbb{E}[\tau|\tau \geq t] = \exp(\mu + \frac{1}{2}\sigma^2)\left(\frac{\Phi\left(\frac{\mu+\sigma^2-\log t}{\sigma}\right)}{1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right)}\right).$$

Substituting in, therefore we have that

$$\mathbb{E}\left[\eta|\rho \geq z\right] = \gamma_\eta \exp\left(\beta_\eta\mu_\rho + \frac{1}{2}\beta_\eta^2\sigma_\rho^2\right) \times \left(\frac{\Phi\left(\frac{\mu_\rho + \beta_\eta\sigma_\rho^2 - \log z}{\sigma_\rho}\right)}{1 - \Phi\left(\frac{\log z - \mu_\rho}{\sigma_\rho}\right)}\right),$$

where γ_η, β_η are as defined above. Finally, we note that

$$\beta_\theta - 1 = \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2} - 1 = \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta - (\sigma_\theta^2 + \sigma_\eta^2 - 2\xi\sigma_\theta\sigma_\eta)}{\sigma_\rho^2} = \beta_\eta.$$

We thus just denote $\beta = \beta_\theta$. Furthermore, we can simplify the expressions for $\gamma_\theta, \gamma_\eta$ by plugging in β to

$$\gamma_\theta = \exp\left(\mu_\theta - \beta\mu_\rho + \frac{1}{2}\left(\sigma_\theta^2 - \beta^2\sigma_\rho^2\right)\right),$$
$$\gamma_\eta = \exp\left(-\sigma_\eta^2/2 - (\beta-1)\mu_\rho + \frac{1}{2}\left(\sigma_\eta^2 - (\beta-1)^2\sigma_\rho^2\right)\right),$$
$$\beta = \frac{\sigma_\theta^2 - \xi\sigma_\theta\sigma_\eta}{\sigma_\rho^2}.$$

Noting that $\mathbb{E}[\theta] = \exp(\mu_\theta + \frac{1}{2}\sigma_\theta^2)$ and $\mathbb{E}\left[\rho^\beta\right] = \exp\left(\beta\mu_\rho + \beta^2\frac{1}{2}\sigma_\rho^2\right)$ the above terms simplify to

$$\gamma_\theta = \mathbb{E}[\theta]/\mathbb{E}\left[\rho^\beta\right],$$
$$\gamma_\eta = \exp\left(-(\beta-1)\mu_\rho - (\beta-1)^2\frac{1}{2}\sigma_\rho^2\right).$$

This implies that $\gamma_\eta \exp\left(\beta_\eta\mu_\rho + \frac{1}{2}\beta_\eta^2\sigma_\rho^2\right) = 1$ and hence

$$\mathbb{E}[\eta|\rho \geq z] = \left(\frac{\Phi\left(\frac{\mu_\rho + \beta_\eta\sigma_\rho^2 - \log z}{\sigma_\rho}\right)}{1 - \Phi\left(\frac{\log z - \mu_\rho}{\sigma_\rho}\right)}\right) = \left(\frac{1 - \Phi\left(\frac{\log\left(ze^{-\beta_\eta\sigma_\rho^2}\right) - \mu_\rho}{\sigma_\rho}\right)}{1 - \Phi\left(\frac{\log z - \mu_\rho}{\sigma_\rho}\right)}\right),$$
$$= \frac{1 - H\left(ze^{-(\beta-1)\sigma_\rho^2}\right)}{1 - H(z)}.$$

Plugging these into (7) we have (18) as desired.

To see the latter part of the Proposition, we will need the following two lemmata.

LEMMA 2. *For any a, $(\Phi_{\mu,\sigma}(z+a) - \Phi_{\mu,\sigma}(z))/\varphi_{\mu,\sigma}(z)$ is decreasing in z.*

PROOF. Note that

$$\frac{\Phi_{\mu,\sigma}(z+a) - \Phi_{\mu,\sigma}(z))}{\varphi_{\mu,\sigma}(z)} = \frac{\Phi\left(\frac{z+a-\mu}{\sigma}\right) - \Phi\left(\frac{z-\mu}{\sigma}\right))}{\frac{1}{\sigma}\varphi\left(\frac{z-\mu}{\sigma}\right)},$$

so by redefining $\hat{z} = \frac{z-\mu}{\sigma}$ it suffices to prove the result for the standard normal distribution. Note that

$$\left(\frac{\Phi(z+a) - \Phi(z)}{\varphi(z)}\right)' = \frac{\varphi(z+a) - \varphi(z)}{\varphi(z)} - \frac{\varphi'(z)}{\varphi(z)}\frac{\Phi(z+a) - \Phi(z)}{\varphi(z)},$$

and therefore, since $\varphi'(z) = -z\varphi(z)$, we have that

$$\left(\frac{\Phi(z+a)-\Phi(z)}{\varphi(z)}\right)' = \frac{\varphi(z+a)-\varphi(z)}{\varphi(z)} + z\frac{\Phi(z+a)-\Phi(z)}{\varphi(z)}.. \tag{26}$$

We note that for $z \geq 0$ and $a \geq 0$

$$\Phi(z+a)-\Phi(z) = \int_z^{z+a} \varphi(s)ds \leq \int_z^{z+a} \frac{s}{z}\varphi(s)ds = -\frac{1}{z}\int_z^{z+a} \varphi'(s)ds$$
$$= -\frac{\varphi(z+a)-\varphi(z)}{z}.$$

For $z < 0$ and $a \geq 0$ we have that

$$\begin{aligned}\Phi(z+a)-\Phi(z) &= \int_z^{\min\{z+a,0\}} \varphi(s)ds + \int_{\min\{z+a,0\}}^{z+a} \varphi(s)ds \\ &\geq \int_z^{\min\{z+a,0\}} \frac{s}{z}\varphi(s)ds + \int_{\min\{z+a,0\}}^{z+a} \frac{s}{z}\varphi(s)ds \\ &= \int_z^{z+a} \frac{s}{z}\varphi(s)ds = -\frac{1}{z}\int_z^{z+a} \varphi'(s)ds = -\frac{\varphi(z+a)-\varphi(z)}{z}.\end{aligned}$$

Plugging into (26) yields that for $a \geq 0$

$$\left(\frac{\Phi(z+a)-\Phi(z)}{\varphi(z)}\right)' \leq 0.$$

To see the case $a < 0$, note that

$$\frac{\Phi(z+a)-\Phi(z)}{\varphi(z)} = \frac{\Phi(-z)-\Phi(-z-a)}{\varphi(-z)} = -\frac{\Phi(-z-a)-\Phi(-z)}{\varphi(-z)}.$$

However, we have already shown that $((\Phi(-z-a)-\Phi(-z))/\varphi(-z))' \geq 0$, and therefore we have the desired result for $a < 0$ also. ■

LEMMA 3. *Consider the lognormal case. We have that $z \mapsto J(z)/z$ is non-decreasing and $\bar{J}(\rho) = \max\{0, J(\rho)\}$ for $\alpha = 0$.*

PROOF. As $J(0) = 0$ we need to show that J is non-decreasing whenever it is positive. As $H(z) = \Phi\left(\frac{\log\rho-\mu_\rho}{\sigma_\rho}\right) = \Phi_{\mu_\rho,\sigma_\rho}(\log\rho)$ we get that the virtual value is given by

$$J(z) = z\left(1 - \frac{\Phi_{\mu_\rho,\sigma_\rho}\left(\log(z)+(1-\beta)\sigma_\rho^2\right) - \Phi_{\mu_\rho,\sigma_\rho}\left(\log(z)\right)}{\varphi_{\mu_\rho,\sigma_\rho}\left(\log(z)\right)}\right).$$

The function $z \mapsto J(z)/z$ is thus non-decreasing whenever positive if the term

$$\frac{\Phi_{\mu_\rho,\sigma_\rho}\left(y+(1-\beta)\sigma_\rho^2\right) - \Phi_{\mu_\rho,\sigma_\rho}\left(y\right)}{\Phi'_{\mu_\rho,\sigma_\rho}\left(y\right)}$$

is decreasing in $y = \log(z)$. By Lemma 2 this is the case. ■

From Corollary 3, recall that the marginal taxes are increasing (respectively, decreasing) if $J(z)/z$ is decreasing (respectively, increasing). In the case of narrow allocative efficiency objective ($\alpha = 1$), this equals

$$\frac{\mathbb{E}\left[\theta\right]}{\mathbb{E}\left[\rho^{\beta}\right]} z^{\beta-1} .$$

Consequently, the marginal tax is increasing for $\beta \in (0, 1)$ and decreasing for $\beta > 1$. In the case of utilitarian objective ($\alpha = 0$), follows from Lemma 3. ■

PROPOSITION 9 (Optimal Mechanism in the Lognormal Case). *Suppose (θ, η) are jointly lognormally distributed and v is strictly concave. Consider a planner concerned with narrow allocative efficiency, i.e., $\alpha = 1$.*

(i) $\beta < 0$: there exists a threshold $\bar{c} > \mathbb{E}\left[\theta\right]$ such that if $\mathbb{E}\left[c\right] \le \bar{c}$, the optimal mechanism is a consumption quota and if $\mathbb{E}\left[c\right] > \bar{c}$, the optimal mechanism is prohibition.

(ii) $\beta > 0$: the optimal mechanism is a consumption tax with strictly increasing marginal tax if $\beta \in (0, 1)$ and strictly decreasing marginal tax if $\beta > 1$.

Consider a utilitarian planner, i.e., $\alpha = 0$.

(iii) The optimal mechanism is a consumption tax with strictly decreasing marginal tax.

PROOF. To see this, observe that when $\alpha = 1$ and β is negative, then $J(z)$ is a strictly decreasing function of z. Therefore the ironed $\bar{J}(\cdot)$ will be a constant, $\bar{J}$. If $c > \bar{J}$ then the optimal solution for the planner is complete shutdown. Conversely, if $c < \bar{J}$ then the solution is a uniform quota of $v'^{-1}(c/\bar{J})$ to all agents. ■

COROLLARY 6. *Suppose (θ, η) are jointly lognormally distributed and $\beta > 1$. Then for a utilitarian planner, i.e. $\alpha = 0$,*

(i) For $\mathbb{E}\left[c\right]$ small enough, the optimal mechanism has uniformly higher consumption than under Laissez-faire.

(ii) For β close to 1, the optimal mechanism will have lower consumption than under Laissez-Faire but higher consumption than under Pigouvian taxation.

PROOF. To see this recall that under Laissez-Faire, an agent with willingness to pay $\rho = z$ will consume $v'^{-1}(\kappa/z)$, whereas under the optimal mechanism for a planner with utilitarian objective, the agent will consume $v'^{-1}((\kappa + \mathbb{E}\left[c\right])/J(z))$. Since $J(z) > z$ when $\beta > 1$ by observation, we have that if c is small enough, the latter will be uniformly higher.

To see the second part, recall that under Pigouvian taxation, an agent with willingess to pay $\rho = z$ will consume $v'^{-1}((\kappa + c\mathbb{E}[1/\eta])/z)$. For β close enough to 1, $J(z) \approx z$ and again the result obtains. ■

PROPOSITION 10. *A mechanism (x, t) maximizes (20) subject to (B) and (IC) if and only if it maximizes (OBJ) for $\alpha = 0$ subject to (B) and (IC) with respect to the type distribution F obtained by setting*

$$\theta = \rho_\theta \, w(\rho_\theta, i, \rho_c)$$
$$\eta = w(\rho_\theta, i, \rho_c)$$
$$c = \rho_c \, w(\rho_\theta, i, \rho_c)$$

where $w(\rho_\theta, i, \rho_c) = \varphi'(U(\rho_\theta, i, \rho_c; x, t))$.

PROOF. Let $\mathcal{U}$ be the set of implementable type-dependent utility functions. As the set of incentive compatible and budget feasible mechanisms is convex and the utility is linear in x and t it follows that $\mathcal{U}$ is convex. We can thus restate the optimization problem of the designer as

$$\max_{u \in \mathcal{U}} \int \varphi(u(\rho_\theta, i, \rho_c)) dF \,. \tag{27}$$

As φ is strictly concave it follows the maximum is unique (if it exists). Let (x^*, t^*) be an optimal mechanism, $u^*(\rho_\theta, i, \rho_c) = U(\rho_\theta, i, \rho_c; x^*, t^*)$ be an optimal utility and (x, t) be another incentive compatible and budget feasible mechanism. As (x^*, t^*) is optimal and $\mathcal{U}$ is convex and φ differentiable we have that

$$0 \leq \int \varphi'(U(\rho_\theta, i, \rho_c; x^*, t^*)) \left[U(\rho_\theta, i, \rho_c; x^*, t^*) - U(\rho_\theta, i, \rho_c; x, t) \right] dF \,. \tag{28}$$

This implies that the optimal mechanism is optimal in (OBJ) with welfare weights $w(\rho_\theta, i, \rho_c) = \varphi'(U(\rho_\theta, i, \rho_c; x^*, t^*))$.

Conversely, we have that any utility $u^*(\cdot) \equiv U(\cdot; x^*, t^*)$ satisfying (28) constitutes a local maximum in (27). As the objective function in (28) is strictly concave over $\mathcal{U}$ any such local maximum is also a global maximum. ■

Finally, we present the results of our calibration when the Pigouvian tax would be \$ 25, rather than \$ 50 as assumed in the main test.

	Avg Trips	Rebate	Flight Taxes					
			1st	2nd	3rd	4th	5th	6th
Allocative	3.75	965.11	96.15	138.2	155.12	166.36	174.86	181.71
Utilitarian	4.51	707.11	185.41	177.10	174.51	172.88	171.68	170.73
Pigouvian	5.96	149.11	25	25	25	25	25	25

TABLE 2. Table summarizing the average number of round trips, average redistribution to all agents and marginal tax of 1st thru 6th flights under Pigouvian, Allocative Efficient and Utilitarian Optimal Taxation under a carbon price of $50/ton.