An Ethical Pollution Tax

Garth Heutel

Abstract

Through theory and numerical simulations, I explore how a pollution tax can be designed to accommodate ethical objections. I consider four types of objections that have been raised by ethicists – based on equity concerns, incommensurability, civic responsibility, and commodification. In each case I show theoretically how a tax can be designed in response, and quantitatively, using a general equilibrium model calibrated to the U.S. economy and carbon dioxide pollution, how these accommodations affect the magnitude of the tax and the resulting emissions level. Some accommodations likely have small effects on the carbon tax rate, while others can cause the tax rate to increase fivefold.

February 2026

Keywords: Pollution pricing; Carbon tax; Philosophy of economics; Ethics

JEL codes: Q58; H23; D63; A13

*Department of Economics, Georgia State University, and National Bureau of Economic Research. Email: gheutel@gsu.edu. Thanks to Andrew J. Cohen, Teevrat Garg, J. Paul Kelleher, Linus Mattauch, Yongsheng Xu, and seminar participants for helpful comments and to Tejendra Pratap Singh for valuable research assistance.

I. Introduction

Pricing pollution via a tax or a tradable permit system is nearly universally advocated by economists and policymakers steeped in the neoclassical economic tradition. The justification, dating back to Pigou, is based on the efficiency of internalizing the externality associated with pollution; without the pollution price, this externality is not reflected in market transactions. While the intuition behind the theoretical justification is clear (albeit subject to numerous simplifying assumptions that may complicate the justification in the real world), the theory ignores many ethical issues that philosophers and economists have struggled with. Many people have argued that a pollution tax, while yielding a cost-effective and efficient abatement strategy in theory, violates ethical norms or is ethically questionable. For example, some argue that a pollution pricing system unethically allows polluters to avoid their civic responsibility to protect the environment simply by paying enough money.

The literature on environmental ethics, including the ethics of environmental policy and pollution pricing, is large, although it is unfamiliar to many environmental economists. Several papers summarize and synthesize this literature, including Caney and Hepburn (2011), Aldred (2012), and Mintz-Woo (2024). As described in Heutel and Kelleher (2024), objections to pollution pricing can be related to its use in designing policy and setting pollution goals, or they can be related to the use of pollution pricing itself as a way of meeting a predetermined goal. At any stage of the policy design process, ethical issues with the use of pricing pollution can arise, and these issues are mostly disregarded in environmental economics.

The purpose of this paper is to investigate whether and how a pollution tax can be designed to address and accommodate these ethical objections. This may appear impossible – any objection to a pollution tax cannot be accommodated by a pollution tax. In fact, I develop a simple stylized model of pollution taxation in which these ethical objections can be incorporated, and I demonstrate how the design of the tax can or cannot address each objection. I consider four types of ethical objections to pollution pricing – based on equity, incommensurability between environmental and market goods, a civic responsibility to protect the environment that cannot be discharged by paying a tax, and the pollution tax's commodification of environmental amenities that ought not to be treated as a commodity. I present theoretical expressions for a pollution tax accommodating each ethical concern, modifying the standard Pigouvian tax equal to marginal external damages. I also present calibrated numerical simulation results to quantify the effects that these ethical accommodations have on the magnitude of the pollution tax and the resulting emissions, though with the proviso that the relatively simple model ignores many real-world complications. Though the model is framed in terms of a pollution tax, the findings generalize to tradable pollution permit systems and other forms of pollution pricing.

While many previous studies have discussed ethical objections to pricing pollution, to the best of my knowledge this is the first study to explicitly incorporate ethical objections to pollution pricing into a neoclassical economic model of pollution taxation and the first economic model that seriously engages with this literature in environmental ethics in this way. Some other studies have incorporated ethical issues into economic models of

climate or environmental policy. For instance, Howarth (2001) presents an overlapping generations model of climate policy that compares sustainability-based and libertarian-based optimization criteria to an efficiency criterion, and Llavador et al. (2011) develop an intergenerational model and optimize climate policy based on maintaining sustainability or an exogenous sustainable growth rate. These papers seriously address alternative ethical models of optimal policy criteria in an economic framework, but they do not attempt to address ethical objections to pollution pricing itself, which is the focus of this study.[1] While many studies in environmental economics have considered distributional and equity implications of pollution pricing (e.g., West and Williams 2004), which is one of the ethical issues that I confront, the other three ethical issues have not been addressed in the economics literature. This paper is very much in the tradition of the "optimal tax" literature (Mirrlees 1976), and specifically the literature on optimal environmental taxes (Bovenberg and de Mooij 1994; Parry 1995; Metcalf 2003). While many extensions to the standard Pigouvian model of pollution pricing have been considered (e.g. other market failures like information spillovers (Jaffe et al. 2005) or market power (Fowlie 2010)), ethical concerns have not.

I find that a pollution tax can be modified in relatively straightforward ways to accommodate these four sets of ethical objections. Objections based on distributional or equity concerns can be accommodated via a tax that is sensitive to distribution, for

[1] Howarth (2001) does report the carbon tax rates that would be required to achieve the optimum allocation under the various ethical criteria. Using the terminology from Goodin (1994) and Heutel and Kelleher (2024), these other studies focus on questions of *policy choice*, while this paper focuses on questions of *policy enforcement*.

instance one that maximizes a weighted utilitarian objective function or a Rawlsian maximin objective function; this result is not novel. Objections based on the incommensurability of environmental goods with market goods can also be accommodated, so long as the tax is designed in such a way that does not impose a commensurability where it is not needed (i.e. the tax can be consistent with value pluralism). A pollution tax can accommodate civic responsibility, so long as it is designed to satisfy a constraint ensuring that all agents' responsibilities are met. If the objection based on commodification boils down to a practical concern with crowding out intrinsic motivations for protecting the environment, then the tax rate can be modified accordingly. Quantitatively, I find that some of these accommodations have only modest effects on the magnitude of the pollution tax, while others can have large effects. For example, under reasonable parameter values in the social welfare function accounting for equity concerns, the pollution tax is about 20%-25% higher than its Pigouvian level (which ignores distributional concerns). However, under a reasonable quantification of a civic responsibility constraint applying to all consumers, the pollution tax can be about five times higher than its Pigouvian level (which ignores such a constraint). These results (both theoretical and quantitative) are meant to be starting points to guide economists and policymakers on designing pollution pricing policy with ethical concerns in mind, rather than providing definitive answers to this complex literature.

In the following section, I briefly summarize four ethical objections to pollution pricing that have been identified by ethicists and philosophers. Then, in section III, I

present the framework for the basic Pigouvian model justifying pollution pricing. Section IV provides the paper's main theoretical and simulation results.

II. Background

Several studies have considered ethical implications of pollution pricing and identified specific ethical issues related to pollution pricing. Summaries of these ethical issues can be found in Caney and Hepburn (2011) or in Heutel and Kelleher (2024). Here, I will briefly summarize four categories of ethical issues – based on equity, incommensurability, civic responsibility, and commodification – to give the reader sufficient background to understand how I will then theoretically and numerically address them in the following sections. Appendix section A.I presents additional details and a more thorough literature review.

II.a. Equity

The standard neoclassical justification for pollution pricing is based on either efficiency or cost-effectiveness criteria (i.e., any pollution price will yield a cost-effective allocation of pollution abatement, and a pollution price set to marginal external damages will yield the efficient level of pollution abatement). Concerns over equity are absent. This has led many to object to pollution pricing on ethical grounds based on distributional issues with the tax. In fact, of all the ethical objections to a carbon cap-and-trade market discussed by Caney and Hepburn (2011), the only one that they find justifiable is the one based on distributional concerns. Standard neoclassical theory might respond by arguing

that the goal of a pollution tax is to target the pollution externality, and that distributional concerns should be addressed through other measures, like lump-sum transfers (appealing to the second fundamental theorem). This argument is unconvincing since in fact pollution taxes ought to be evaluated on their own, and their distributional impacts should be important considerations in their evaluation.

II.b. Incommensurability

The justification for a pollution tax relies on comparing costs and benefits of pollution abatement with other costs and benefits. That is, the benefits of reduced pollution – cleaner air, improved health, reduced global warming – are directly compared in magnitude to the monetary costs of reduced pollution – abatement technology costs faced by firms, etc. The derivation of a Pigouvian tax requires that all costs and benefits are *commensurable*, in the same way that the costs of various market goods are commensurable when consumers make consumption decisions. But many philosophers say that not all "goods" are or should be commensurable with each other. Environmental goods like clean air or biodiversity may be incommensurable with market goods, and the trade-offs that are inherently made in justifying pollution pricing may be morally inappropriate (Anderson 1995).

II.c. Civic Responsibility

Under pollution pricing, polluters can pollute as much as they want, so long as they pay the per-unit price on pollution. This strikes some ethicists as morally problematic,

since it implies that polluters can escape from a collective or civic responsibility simply by paying their way out (Sandel 2012). If it is the case that there are some moral obligations that individuals face, for example, not polluting the shared environmental excessively, then buying one's way out of these obligations is not morally permissible, even if it may yield an economically efficient or cost-effective outcome.

II.d. Commodification

Placing a price on pollution strikes some philosophers as equivalent to treating clean air or clean water as a market good. This form of commodification may be seen as morally objectionable. Satz (2010) provides many examples of what she calls "noxious markets," or types of goods or services that should not be provided by markets. She cites child labor as an example. Others may argue that markets in pollution fall under that same objection – the price that a pollution tax or cap-and-trade system places on pollution improperly commodifies something that should not be treated as a market good. Page (2012) makes this argument, claiming that the commodification created by these policies can erode normative ideals of political legitimacy and procedural justice.

III. Framework

The model is a very basic and standard optimal tax model in an economy with a negative externality, providing a framework for justifying the Pigouvian claim that pollution pricing yields an optimal outcome. The model is derived from the model in Phaneuf and Requate (2016, Ch. 1).

There are N consumers. Each consumer has a utility function over two consumption goods, x and z, and the level of the pollution externality, E. Consumer i's utility function is $U_i(x_i, z_i, E)$, where x_i and z_i are consumer i's consumption of each of the two goods, and the pollution externality E is not indexed by i since all consumers are exposed to the same level of it.[2] The utility function allows for general elasticities of substitution between the goods. There is a fixed supply of labor l in the economy that is distributed across production in the two sectors: $l_x + l_z = l$; there is no capital in the model and no unemployment. Good x is the "dirty" or "polluting" good since its production involves the creation of pollution. I model this by setting pollution E as a factor of production for the dirty good as is common in the literature. The dirty good production function is $f(l_x, E)$, so that the dirty good resource constraint is $f(l_x, E) = \sum_{i=1}^{N} x_i$. The clean good production function is a function of just the labor in the clean sector l_z, so the clean good resource constraint is $g(l_z) = \sum_{i=1}^{N} z_i$.

III.a. Social Planner's Problem

I first consider the social planner's problem, and I define the solution to this problem as the "optimal" outcome. The social planner maximizes a social welfare function that is a weighted sum of each consumer i's utility level, where the weight on consumer i's utility is γ_i. That is, the social welfare function (the planner's maximand) is $\sum_{i=1}^{N} \gamma_i U_i(x_i, z_i, E)$. The social planner's problem is thus

[2] There is only the one, global pollutant E in the model. Co-pollutants, including local pollutants that could yield differential exposure to different households (E_i) are not modeled.

$$\max_{\substack{\{x_i\},\{z_i\},\\ l_x,l_z,E}} \sum_{i=1}^{N} \gamma_i U_i(x_i, z_i, E)$$

such that

$$f(l_x, E) = \sum_{i=1}^{N} x_i$$

$$g(l_z) = \sum_{i=1}^{N} z_i$$

$$l_x + l_z = l$$

The problem can be solved with the Lagrangian method, yielding first-order conditions involving the endogenous choice variables and Lagrangian multipliers which define the optimal outcome.

This approach explicitly imposes a welfarist conception of optimal policy, by using a social welfare function based on individual utility functions. In section IV of this paper when I modify this social welfare function to accommodate ethical objections to pollution pricing, I will be maintaining this welfarist approach. There are many philosophical objections to welfarism and to utilitarianism (e.g. Smart and Williams 1973); I will not be engaging those objections. The ethical objections to pollution pricing that I engage with are unrelated to ethical objections to welfarism or utilitarianism.

Furthermore, a social-welfare-maximizing outcome can always be equivalent to a Pareto-efficient outcome, under certain conditions, given proper weights in the social welfare function.[3] Pareto efficiency does not require cardinal, interpersonally comparable

[3] See Proposition 16.E.2 in Mas-Colell et al. (1995).

utility functions, as is required for a social welfare function approach. A long tradition in normative economics attempts to avoid those assumptions on utility functions by considering Pareto efficiency as their optimizing framework. By contract, I am explicitly invoking a social welfare function approach and therefore relying on cardinal and interpersonally comparable utility functions. Ethical or conceptual issues with cardinality or interpersonal comparability of utility functions are not addressed here. See Kelleher (2025) for more discussion of this issue in the context of environmental policy and the social cost of carbon.

III.b. Decentralized Solution

Next, I consider a decentralized model. There is a competitive economy, with output good prices p_x and p_z, a price for the labor input w, and an income level y_i for each consumer i. Prices and income are endogenously determined in general equilibrium. There is a representative price-taking firm producing x and another producing z. There is also an emissions tax τ that the dirty good producing firm must pay for each unit of emissions that it produces. Emissions tax revenues are returned to consumers as lump-sum transfers.[4] Each consumer i is assigned an ownership share of the profits of each firm (θ_{ix} and θ_{iz}), the labor income (θ_{il}), and the emission tax revenues ($\theta_{i\tau}$). Consumer i maximizes utility subject to their budget constraint:

$$\max_{x_i, z_i} U_i(x_i, z_i, E)$$

[4] In Howarth (2001), the only revenues available for lump-sum redistribution are the pollution tax revenues, but he also models intergenerational payments to compensate for environmental quality changes. My model is static and does not consider intergenerational issues.

such that

$$p_x x_i + p_z z_i = y_i$$

Consumer i's income y_i is determined by the ownership shares and total economy-wide income:

$$y_i = \theta_{ix}\pi_x + \theta_{iz}\pi_z + \theta_{il} wl + \theta_{i\tau}\tau E$$

Here $\pi_x = p_x f(l_x, E) - wl_x - \tau E$ is firm x's profit, and $\pi_z = p_z g(l_z) - wl_z$ is firm z's profit. Total labor income is wl since the total labor supplied (inelastically) is l, and τE is total emissions tax revenue.

The dirty good producing firm's profit maximization problem is

$$\max_{l_x, E} p_x f(l_x, E) - wl_x - \tau E$$

The clean good producing firm's profit maximization problem is

$$\max_{l_z} p_z g(l_z) - wl_z$$

Each consumer's utility maximization problem and both firms' profit maximization problems can be solved with the Lagrangian method, generating first-order conditions. These first-order conditions along with the resource constraint equations define the decentralized competitive equilibrium with an exogenous pollution tax. Prices are determined in equilibrium; I normalize $p_x = 1$.

While I omit here the first-order conditions for brevity, it can be shown (see Phaneuf and Requate 2016, p. 11) that the optimal solution (again, "optimal" here is defined as the solution to the planner's problem described above) is reached when the pollution tax τ is set to $\tau = -p_x \sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i}$. This is the Pigouvian tax, setting the tax rate to equal the monetized value of marginal external damages, internalizing the pollution externality to the

polluting firm. Importantly, this optimal Pigouvian tax depends on the availability of lump-sum transfers, beyond just the lump-sum return of pollution tax revenues. As shown below, if the pollution tax is the only instrument available, and lump-sum transfers (other than the return of the pollution tax revenue) are not available, then the social-welfare-maximizing tax rate does not necessarily equal this Pigouvian tax rate, depending on the distribution of income across the economy.

III.c. Regulator's Problem

Finally, I consider what I call the regulator's problem, which is to choose a pollution tax rate τ to maximize social welfare, subject to the constraints of the decentralized model, including the exogenous ownership shares. That is, the regulator chooses τ to maximize $\sum_{i=1}^{N} \gamma_i U_i(x_i, z_i, E)$, subject to the resulting solution being the solution of the decentralized model described above. (As opposed to the social planner, who can directly choose all the consumption levels of each consumer without regard for the outcome being a decentralized equilibrium). The regulator returns pollution tax revenue through lump-sum transfers based on each individual's share $\theta_{i\tau}$, but does not have the ability to make any other lump-sum transfers. Because of this limitation, the tax rate chosen by the regulator is not necessarily the Pigouvian tax rate. If the regulator had access to any level of lump-sum transfers, then the regulator would use those transfers to achieve a given distribution of income and always apply the Pigouvian tax rate to pollution (consistent with the second fundamental welfare theorem). In the simulations below, I show how the regulator's

chosen tax rate can vary with a number of parameters and generally is not equal to the Pigouvian tax rate.

III.d. Functional Form and Calibration

The theoretical results presented below do not depend on functional form or parameter values. But I complement the theoretical results with calibrated numerical simulations, to add additional intuition behind the theory's results and to quantify. The simulation is a relatively simple general equilibrium economy with two representative firms and a few representative households; it is not a CGE model or an IAM. The quantitative results are not intended to be definitive values for optimal policy, but rather to guide the reader towards understanding the theory's implications for policy design.

I use the following functional forms in the calibrated simulations. Each consumer's utility function is separable in both consumption goods and linear in environmental quality: $U_i = x_i^{\alpha} + z_i^{\omega} - \beta E$. The separability assumption is solely for tractability, allowing me to analytically solve the equilibrium using first-order conditions. The dirty good production function is Cobb-Douglas: $f(l_x, E) = l_x^{\kappa} E^{1-\kappa}$. The clean good production function is a power function: $g(l_z) = l_z^{\mu}$. These functional form assumptions mean that the dirty good sector features constant returns to scale production but the clean good sector does not (unless $\mu = 1$).

Details on the calibration are provided in Appendix section A.II, and Appendix Table A1 lists the parameter values. Here I summarize. I set $N = 5$, allowing each representative household to represent one income quintile of the U.S population. Income shares are

based on data reported in CBO (2008), which features significant income inequality: the richest income quintile earns 87% of firm profits and 50% of labor income. Pollution damages and the Pigouvian tax rate are calibrated based on the DICE-2023 model (reported in Barrage and Nordhaus 2024) and Paul et al. (2013). The calibrated values are such that the efficient Pigouvian carbon tax (ignoring the ethical concerns that I will introduce) is exactly \$125 per ton of CO_2. The efficient level of pollution is about 33% lower than the business-as-usual (BAU) level.

IV. Results

I present both theoretical and quantitative simulation results. I consider the implications of the four ethical concerns that I discussed earlier: equity, incommensurability, civic responsibility, and commodification. The theoretical results identify how an optimal tax is modified to accommodate the various ethical concerns. The quantitative simulation results explore the magnitude of these effects on the tax rate and on the resulting quantity of emissions. I consider each of these four ethical considerations independently, one at a time, in the following four subsections.

IV.a. Equity

I begin by exploring how equity concerns affect environmental tax design under the weighted utilitarian social welfare function $\sum_{i=1}^{N} \gamma_i U_i(x_i, z_i, E)$, where the social welfare weights are the γ_i values. An unweighted Benthamite social welfare functions has all of these γ_i values equal to each other, though other weights are possible. Argenziana and

Gilboa (2019) provide a theoretical justification for weighted utilitarian social welfare functions based on moral intuitions.[5] I will vary the variance of these weighting values to explore the implications for the pollution tax that solves the regulator's problem. As described earlier, the Pigouvian tax ensures the optimal level of pollution only when policy also includes lump-sum transfers across households. This Pigouvian tax will then be invariant to the distribution of income across households, though the levels of the lump-sum transfers will depend on this distribution. It is more interesting to consider the regulator's problem, in which a pollution tax can be chosen to maximize social welfare but lump-sum transfers are unavailable (other than the return of pollution tax revenue). In Figure 1 I plot simulation results showing how the emissions tax that solves the regulator's problem varies with the variance in the vector γ describing the distribution of welfare weights. I start with the unweighted case where all the weights are equal to one, and thus the variance is zero (corresponding to zero on the x-axis of Figure 1). I then increase the variance by increasing the weights γ_i on the poorest households and decreasing the weights γ_i on the richest households so that all five weights still sum to five and so that there is an equal spacing between the five values. I continue this until the weight for the poorest household, γ_1, equals 2.5, and the weight for the richest household, γ_5, equals

[5] Weighting of individual utility functions in a social welfare function is not identical to weighting consumption or income (or changes in consumption or income) in cost-benefit analysis, as in Cowell and Gardiner (1999). Those welfare weights reflect heterogeneous marginal utility arising from unweighted utilitarian welfare functions, and typically lower-income households are given a higher social welfare weight, reflecting the diminishing marginal utility of wealth (Adler 2016).

0.25.[6] For each set of weights, I solve the regulator's problem, where the distribution of wealth (given by the θ parameters) is unchanged.

Figure 1 shows that a higher variance in γ, corresponding to a higher degree of aversion to income inequality, yields a higher emissions tax. When the regulator is unconcerned with inequality (variance of zero), the tax rate is about $121 per ton carbon dioxide. This is notably slightly lower than the Pigouvian tax, which is exactly $125 per ton (the damage parameter is calibrated to reach this optimal Pigouvian tax). Even without a concern for inequality in the social welfare function, the regulator sets an emissions tax lower than the Pigouvian level because of the income inequality present in the economy. Because the regulator does not have access to a lump-sum transfer, the inequality in income affects what tax it would prefer to levy. When the variance of γ is at its highest level in Figure 1, which again roughly corresponds to weighting based on the inverse of relative income, the emissions tax is about $150 per ton, 20% higher than the Pigouvian level.[7]

An alternative way to handle distributional concerns is to impose a social welfare function based on a Rawlsian concept of justice rather than a weighted utilitarian function. As discussed in the Appendix, it is debatable as to whether imposing a social welfare function is in fact faithful to Rawls, though it is often used. I explore the implications here by replacing the weighted utilitarian social welfare function with an isoelastic social welfare function:

---

[6] This set of values for the distributional weights is roughly equivalent to weighting households inversely proportional to income, as would be done weighting income changes in a cost-benefit analysis with logarithmic utility of consumption (Cowell and Gardiner, 1999), though see the previous footnote.

[7] These simulations, and all simulations in the main text of this paper, keep the exogenous income distribution (described by the θ parameters) fixed. In Appendix section A.III, I present simulations that vary this distribution.

$$\frac{1}{1-\eta}\sum_{i=1}^{N}\left(U_i(x_i, z_i, E)\right)^{1-\eta}$$

The parameter $\eta > 0$ describes aversion to inequality in utility. As $\eta \to \infty$, the isoelastic social welfare function approaches a Rawlsian maximin function.[8]

Theoretically, under isoelastic social welfare, the planner's problem yields the same solution as under unweighted utilitarian social welfare. The planner is unconcerned with the initial distributions of wealth in the economy and chooses a final distribution to minimize inequality given this functional form, which is always an equal distribution. The isoelastic social welfare function becomes interesting when applied to the regulator's problem, for a regulator choosing an emissions tax in an economy with existing inequalities, described by the income share parameters θ.

Figure 2 shows the solution for the emissions tax that solves the regulator's problem under the isoelastic social welfare function as the inequality aversion parameter η is varied from 0.1 to 2.[9] Again the inequality in the economy governed by the income share parameters θ is unchanged. As inequality aversion increases, the emissions tax that solves the regulator's problem increases. Just like with the weighted utilitarian social welfare function, under the isoelastic social welfare function the importance of equity concerns in the choice of the optimal emission tax rate is apparent. Given the underlying inequality in these simulations driven by the exogenous income shares, a greater concern

[8] Jacobs and de Mooij (2015) use a Bergson-Samuelson social welfare function to study environmental policy, which is a generalization of the isoelastic.

[9] Bucholz and Schumacher (2010) report on empirical work attempting to measure an appropriate value for η and find values vary widely, from close to 0 to between 2 and 3. Similarly, Del Campo et al. (2024) review the literature attempting to identify the value for η specifically related to inequality aversion regarding climate policy, and also report a wide range (see their Figure 1). Based on this, here I vary η from 0.1 to 2.

for equity raises the regulator's choice of the tax rate. Over a reasonable range of the η parameter value, given that the empirical literature ranges over these values, the optimal emissions tax varies from about $125 to $145 per ton.[10]

These figures demonstrate how an emissions tax can be designed to accommodate distributional concerns. While the textbook formula for an efficient Pigouvian tax does not explicitly include distributional terms, these distributional concerns are implicitly included in the tax formula and reflected in the calculation of the tax. My motivating example was the concern for equity across income groups, though this model could also be applied to equity across countries (assuming the i is indexing countries rather than income groups). As discussed earlier, this finding is not novel, as many previous authors have explored (often with more complex models than the one here) how equity and distributional concerns affect pollution pricing design. Thus, I now turn to the novel features of my model, which is how the tax can be designed to accommodate three other sets of ethical objections.

IV.b. Incommensurability

If environmental goods are incommensurable with market goods and cannot be combined into a social welfare function, can a pollution tax be justified? Here I propose that the answer is yes. Consider a social planner's problem. Each individual consumer has a utility function over market goods (x_i and z_i). This utility function could also include

[10] Appendix Figures A4 and A5 show how the emissions level varies under the regulator's problem with respect to the variance in the utilitarian social welfare weights or the inequality aversion parameter in the isoelastic social welfare function, respectively.

the level of environmental quality or emissions E, which could reflect either a direct valuation of environmental goods (willingness to pay for clean air) or how the level of environmental quality affects valuation of and willingness to pay for market goods (if some of them are complements or substitutes to environmental quality). If so, one could argue that the incommensurability arguments apply to a social planner's problem – the planner cannot trade off between market and environmental goods – but it still allows for individual consumer preferences to include environmental goods, either directly or indirectly. For the time being I will take this approach and allow the consumer's utility function to include emissions: $U_i(x_i, z_i; E)$.

The social planner has an element of the social welfare function that is a function of the individual consumers' utility levels; allow that to be a weighted utilitarian social welfare function (though adopting a different form does not affect the following arguments or intuition): $\sum_i \gamma_i \, U_i(x_i, z_i; E)$. But, because environmental quality is incommensurable with consumer utility over market goods, the social welfare function must have an additional element. Suppose that this element is simply a scalar-valued function over the level of emissions: $V(E)$, where $V'(E) \leq 0$ indicating that more pollution yields a lower level of this element of social welfare. Thus, the social welfare function is in fact a vector-valued function rather than scalar-valued; incommensurability means we cannot combine through a weighted sum or any other comparison the welfare from market goods with the welfare from the environment. The vector-valued (dimension = 2) social welfare function is:

$$W = \left[\sum_i \gamma_i \, U_i(x_i, z_i; E); V(E) \right]$$

The planner seeks to maximize social welfare. But how can the planner maximize a vector-valued function? Again, the planner cannot simply combine the two elements of social welfare, for example by taking a Euclidean norm, since that would violate incommensurability.

There is a field of mathematics called multi-objective optimization, also called Pareto optimization, that handles cases like this that are attempting to find an optimum for vector-valued objectives.[11] It might be the case that the solution that maximizes the first element of the welfare vector also maximizes the second element, though that is unlikely. Generally in this field, the focus is on finding a series of solutions based on the Pareto frontier (a concept familiar to economists): the set of solutions that cannot be improved for any one element without degrading the outcome for any other element.

The solution concept that I explore here is based on maximizing the part of the vector welfare function based on consumer utility over market goods with the condition that the part of the vector welfare function based on the environment achieves some minimum value. That is, the planner's objective function is

$$\max \sum_i \gamma_i \, U_i(x_i, z_i; E) \text{ such that } V(E) \geq \bar{V}$$

Here $\bar{V}$ is the exogenously-determined minimum value of environmental quality that the planner must satisfy. The origin of this value is not explored here; part of the incommensurability argument is that a value like this cannot be determined by some economic model based on comparing monetary costs and benefits. Presumably society

[11] See Deb (2011) for an introductory text. Janssen (2012) covers how it is used in environmental management problems, including in environmental and resource economics.

reaches the value for $\bar{V}$ through some other process involving collective ethical decision-making, and given this value, the social welfare function approach can be used to find an optimal allocation over the market goods and the environment, respecting the incommensurable constraint on environmental quality. This approach, with a constraint over the minimum level of acceptable environmental quality, is not the only approach that can be taken allowing for incommensurability, though it is a starting point.

Given this specification of the planner's problem, the first-order conditions can be derived. They are identical to the first-order conditions to the standard utilitarian social welfare function maximization without the environmental quality constraint, except for the first-order condition with respect to the level of pollution. Because of that constraint, there is an additional Lagrangian multiplier in the problem; call it λ_V. There are two possible cases for the problem's solution. First, if this multiplier $\lambda_V = 0$, then the constraint does not bind, and the planner's solution is identical here compared to the standard case without the environmental quality constraint. That case represents the case where this environmental quality constraint is already satisfied at the solution to the planner's problem without this constraint. Second, if the multiplier $\lambda_V > 0$, then the constraint binds, so $V(E) = \bar{V}$, and the multiplier λ_V appears in the first-order conditions. These first-order conditions can be manipulated and compared to the first-order conditions to the decentralized problem (which is identical here as in the previous section, since the choice of the social welfare function does not affect the decentralized problem) to show the expression for the optimal pollution tax when the environmental quality constraint binds:

$$\tau = -p_x \left[\sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i} + \frac{\lambda_V}{\lambda_x} V'(E) \right]$$

The first term within the brackets is identical to the standard Pigouvian tax that was derived earlier. It still contains partial derivatives of the consumers' utility functions with respect to pollution; as discussed earlier, here I take the approach that the planner can still take into account how a consumer values the environment relative to market goods, though the planner himself cannot do so which leads to the vector-valued social welfare function. The second term in the above expression for the optimal tax accounts for the environmental quality constraint. The multiplier λ_V is the shadow value of this constraint, divided through by the multiplier on the resource constraint for good x, λ_x, to get it in monetary terms. This optimal tax will generally be higher the more strongly the constraint binds (the higher the value of λ_V, and it will be higher the more steeply that pollution affects environmental valuation ($V'(E)$).

Figures 3 through 5 present simulation results shedding light on the intuition behind this tax consistent with incommensurability. In these simulations I use all the same functional forms and parameter values used in the previous simulations. The new component in these simulations is the environmental valuation function $V(E)$ and the constraint $\bar{V}$. I use a simple negative logarithmic for the function: $V(E) = -\ln(E)$, indicating that the function decreases in pollution E, which is in arbitrary units. The minimum environmental quality constraint value $\bar{V}$ is varied in these simulations to cover ranges where it is and is not binding. Figure 3 presents the optimal pollution tax from the social planner's problem, and Figure 4 presents the optimal level of emissions from the

social planner's problem. These figures demonstrate how the tax is impacted by whether the constraint binds. With a low value of $\bar{V}$, the constraint is not binding, so the optimal tax just equals the Pigouvian tax (exactly $125 per ton) and is unaffected by $\bar{V}$. But once $\bar{V}$ reaches a high enough value, the constraint binds, so that the emissions level and the emissions tax are determined by the environmental constraint rather than by any Pigouvian-style calculus. As $\bar{V}$ gets larger, the resulting optimal pollution tax gets larger, and the optimal emissions level gets smaller.

What is the interpretation of the values of $\bar{V}$? As described earlier, these values must be determined outside of any economic calculus, because of the argument from incommensurability. The units of $\bar{V}$ are not interpretable; rather, as $\bar{V}$ varies we see how the constraint impacts (or doesn't impact) emissions. When $\bar{V}$ is lower than 2.1, the constraint is not binding. At higher values of $\bar{V}$, the constraint is binding, so the interpretation of the value of $\bar{V}$ is that value for which the given emissions level just satisfies the environmental quality constraint. For example, $\bar{V} = 2.2$ corresponds to an environmental quality constraint that is satisfied when the emissions level is about 56% relative to the business-as-usual level.

Figure 5 presents the emissions tax that solves the regulator's problem rather than the social planner's problem. Qualitatively, the relationship between the constraint magnitude and the tax is similar to that for the planner's problem tax. The tax is constant and unaffected by the constraint for low values of the constraint (though in the regulator's problem, this constant tax rate is lower than the Pigouvian level of $125 per ton because of the inequality in the economy), and the tax increases as the constraint gets high enough to

be binding. For a given value of the constraint $\bar{V}$, the regulator's tax is always slightly lower than the planner's tax.[12]

These simulations arise from a social welfare function that respects the incommensurability of market goods and environmental goods. The resulting pollution tax also respects this incommensurability. By taking into account the environmental constraint and not directly comparing environmental and market benefits, the results demonstrate that a pollution tax can indeed be designed to accommodate philosophical concerns over the incommensurability of different types of goods.

There are many caveats from this simple exercise. Notably, the form in which incommensurability is incorporated into a social welfare function is very simple: an additional constraint on environmental quality. Alternate ways to incorporate the value of environmental quality will likely not lead to such straightforward conclusions; Appendix section A.IV. explores one such alternative.

IV.c. Civic Responsibility

Can a pollution tax be designed to accommodate the ethical claim that individuals have a moral or civic responsibility to do their part to protect the environment? The key objection to a pollution tax based on the civic responsibility argument is that paying the tax allows one to evade any potential responsibility, and this responsibility should not be something that you can just buy your way out. However, I demonstrate here that a pollution tax can in fact be designed to accommodate this moral claim.

[12] Appendix Figure A6 presents the resulting emissions levels from the regulator's problem as $\bar{V}$ is varied.

The intuition behind this result is analogous to that from the previous section. If a civic responsibility exists, then this responsibility can be incorporated into the planner's problem through a binding constraint. The planner's problem is to maximize social welfare but respecting the fact that the burden of responsibility is shared in a morally appropriate manner. That can be done through a constraint that mandates each individual meets their civic responsibility.

In this paper's model, pollution is emitted by the single polluting firm. In order to be able to attach civic responsibility over pollution to individuals (consumers), the pollution emitted by the firm must be "matched" to each individual, denoting the amount of total pollution that that individual is responsible for. Given the total amount of pollution E, let E_i represent the amount of pollution that consumer i is responsible for. All pollution is assigned to a single individual, so that $\sum_{i=1}^{N} E_i = E$. There are multiple ways in which pollution can be assigned to individuals. One way is based on consumption of the polluting good x; in this case the pollution shares are defined as $E_i \equiv \frac{Ex_i}{\sum x_i}$. A second way is based on the ownership share in the polluting firm; in this case the pollution shares are defined as $E_i \equiv \theta_{ix} E$. Other definitions of the pollution share are possible as well.[13] Each individual's pollution share E_i is observable by the regulator; issues of monitoring or imperfect observability are ignored (this applies also to the carbon tax itself).

Next, there must be a constraint that each individual faces based on their actual pollution share and how much they are allowed to be responsible for. Let R_i be this value,

[13] For example, in a model with a household-level labor-leisure choice, pollution shares could be based on the labor that each household allocates to the dirty sector.

i.e. the maximum amount of pollution that consumer i be allowed to emit according to the civic responsibility argument. The value could be different for different consumers, perhaps reflecting their ownership shares in the economy or wealth levels. Or it could be a function of total emissions $R_i = R_i(E)$.

The planner's problem becomes

$$\max \sum_i \gamma_i \, U_i(x_i, z_i; E) \text{ such that } E_i \leq R_i \; \forall i$$

The constraints apply to each individual consumer, reflecting the civic responsibility that each consumer has to ensure that their pollution does not exceed the amount they are allowed to emit. In the Lagrangian specification of this constrained optimization problem, the multiplier on the civic responsibility constraint for consumer i is λ_{Ri}. The first-order conditions for this Lagrangian problem can be manipulated and compared to those of the decentralized problem with the polluting firm facing a pollution tax, in the same manner as in the previous subsections, to arrive at an expression for the optimal emission tax that will induce the level of emissions that solves this constrained optimization problem:

$$\tau = p_x \left[-\sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i} + \sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i} \cdot \left(\sum_j \frac{\lambda_{Rj}}{\lambda_x} \cdot \left(\frac{-\partial E_j}{\partial x_i} \right) \right) - \sum_i \frac{\lambda_{Ri}}{\lambda_x} \cdot \left(\frac{\partial R_i}{\partial E} - \frac{\partial E_i}{\partial E} \right) \right]$$

The first term in the brackets is identical to the Pigouvian tax in the standard model without the civic responsibility constraint. The second term reflects how the constraint affects the optimal tax through consumer i's consumption of the polluting good x_i affecting their emissions share E_i. If consumer i's emission share is independent of their consumption of the polluting good x_i, for example if their emissions share is just a function of their polluting firm ownership share θ_{ix}, then this second term drops out. The third term reflects how the

constraint affects the optimal tax through how the total level of emissions E is related to each individual's emissions limit R_i and responsible emissions E_i. These values can be defined in different ways, which may or may not be related to total emissions E, affecting how this term comes into play. The last two terms involve λ_x, the multiplier on the resource constraint for good x.

This general expression for the optimal pollution tax can be simplified under assumptions about the nature of the civic responsibility constraint. For instance, if the responsible emissions share E_i is only a function of total emissions and consumer i's ownership share in the polluting firm, so that $E_i \equiv \theta_{ix}E$, then the second term drops out, and the second part of the difference in the third term ($\frac{\partial E_i}{\partial E}$) is a constant θ_{ix}. And, if the emissions limit R_i is independent of total emissions, then $\frac{\partial R_i}{\partial E} = 0$, simplifying the final term. Under both of these assumptions, the optimal tax is

$$\tau = -p_x \left[\sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i} - \sum_{i} \frac{\lambda_{Ri}}{\lambda_x} \cdot \theta_{ix} \right]$$

An important difference exists between this optimal emissions tax, for the civic responsibility constraint, and the optimal emissions tax for the incommensurability constraint in the previous subsection. Because the incommensurability constraint was over the total amount of emissions, the optimal tax induced a level of emissions guaranteed to satisfy that constraint. But, the civic responsibility constraints apply to the responsible emissions for each individual. The optimal tax merely ensures the optimal overall level of emissions E; it does not ensure that each individual's responsible emissions E_i satisfies the constraint $E_i \leq R_i$.

To see this, Figure 6 plots simulation results for the planner's problem. In these simulations, I impose that each individual household's pollution share E_i is defined based on their consumption of the polluting good, so $E_i \equiv \frac{Ex_i}{\sum x_i}$. I set the civic responsibility constraint R_i to equal 20% of the optimal level of emissions for households $i = 1$ through 4, representing an equal responsibility across each household, and then I vary the constraint that applies to the fifth household R_5 from about 12% to about 35%. Figure 6 shows how the optimal emissions tax from the planner's problem varies with this constraint R_5. In the planner's problem the fact that household 5 happens to be the richest household is irrelevant; the planner ignores the income shares θ_{ix}, etc. When household 5's constraint is at 20% or higher, then none of the household's civic responsibility constraints R_i bind, so the outcome is identical to the planner's problem in the base case without these constraints, yielding an optimal tax of $125 per ton. But once household 5's constraint drops below 20%, its constraint binds. Reducing the maximum amount of pollution that household 5 is allowed to be responsible for increases the value of the optimal emissions tax. When the constraint is lower, optimal emissions is lower, so the optimal emissions tax must be higher to induce it.[14] Importantly, for many parameter values in these simulations, the optimal tax under the planner's problem does not in fact ensure that all consumers satisfy their civic responsibility constraints. For these values, consumer 5's emissions E_5 are greater than their maximum value R_5; this is because consumer 5 has a higher share of ownership wealth in these simulations. Without access

[14] Appendix Figure A7 shows how the resulting optimal emissions level varies with R_5 under the planner's problem.

to lump-sum transfers, a regulator cannot ensure that the civic responsibility constraints are satisfied by implementing the theoretical optimal emissions tax.

This brings us to the regulator's problem, for which the simulations are presented in Figures 7 and 8. The same variation in R_5 is presented here, keeping the constraints R_i fixed at 20% for the other four consumers as in Figure 6, but here under the regulator's rather than the planner's problem. The optimal tax rates from the regulator's problem presented in Figure 7 are not the theoretical optimal tax in the text above. Rather, these tax rates are those that the regulator implements to maximize social welfare while ensuring that all civic responsibility constraints bind. The magnitude of these tax rates is much higher than the Pigouvian value of $125 per ton. For instance, when $R_5 = 20\%$, so that all five consumers face the same civic responsibility constraint, the tax that maximizes social welfare is about $525 per ton, almost four times higher than the Pigouvian value. Because consumer 5 is so much richer than the other consumers (earning 87% of firm profits and 50% of labor income), the regulator needs to impose a much higher tax rate than the Pigouvian level to ensure that this consumer's constraint is satisfied. Figure 8 shows the resulting overall emissions levels that arise from the regulator's problem. They are much lower than the optimal emissions levels from simulations without the civic responsibility constraints; for example, when $R_5 = 20\%$, emissions are about 25% of the business-as-usual level, compared to about 67% of the business-as-usual level without the civic responsibility constraints. The regulator recognizes and responds to the (exogenous) income inequality in the economy and the fact that the regulator cannot address it with lump-sum transfers. The overall emissions level is substantially decreased since that is the most cost-effective

way to reduce the richest consumer's emissions share to a level that satisfies their civic responsibility constraint. Even under a distribution of civic responsibility that appears equitable (all consumers facing the same constraint), the pre-existing inequality in society substantially affects the tax magnitude and the resulting emissions. These simulations demonstrate how tax policy can vary with the magnitude of a civic responsibility constraint, and how this constraint interacts with the existing inequality within an economy.

IV.d. Commodification

If placing a price on pollution improperly treats the environment as a commodity, then it seems tautological that no pollution price can be designed to overcome this ethical objection. That is, the objection is essentialist, and fundamental to the nature of the policy, as described in the appendix. However, another, non-essentialist, interpretation of the commodification objection is the fact that commodifying the environment will crowd out intrinsic motivation for protecting the environment. If the objection based on commodification boils down to objecting on the basis of this crowding out phenomenon, then the prospect of designing a pollution tax to accommodate this objection is valid.

In this section I modify the model so that there exists intrinsic motivation for protecting the environment among some agents in the economy.[15] Pollution is created by the single representative polluting firm, so I impose that the firm's objective function includes some aspect representing intrinsic motivation, rather than intrinsic motivation

[15] See Fehr and Charness (2025) for a review of the literature on social preferences, including intrinsic motivations for contributing to socially-valuable goods.

appearing in the consumers' utility functions. The representative polluting firm in this model is owned by the consumers, so people's intrinsic motivation can be incorporated through the firm's behavior as I do here. The representative polluting firm's profit function, given a pollution tax τ, remains $\pi_x = p_x f(l_x, E) - w l_x - \tau E$. But now the firm no longer acts solely to maximize profit; it also considers its intrinsic motivation for protecting the environment. The firm's objective in this case is to choose its inputs l_x and E so as to maximize the following function: $\pi_x - h(E, \tau) = p_x f(l_x, E) - w l_x - \tau E - h(E, \tau)$. The additional term $h(E, \tau)$ represents the cost arising from intrinsic motivation of the firm. It is subtracted from the profit π_x, meaning that the firm will sacrifice some profit in order to decrease the cost from the intrinsic motivation. The intrinsic motivation term satisfies $\frac{\partial h}{\partial E} > 0$ and $\frac{\partial h}{\partial \tau} < 0$; a higher level of emissions E increases this intrinsic cost, while a higher pollution tax τ reduces it. Both effects reflect a crowding out phenomenon, either crowding out from the tax itself or from the level of emissions the firm is responsible for.[16]

This is how the polluting firm behaves; the non-polluting firm and the consumers behave the same way that they do in the baseline model. Likewise, the social welfare function that the planner maximizes is identical to the social welfare function in the baseline model. The planner does not include the intrinsic motivation term $h(E, \tau)$ in its social welfare function; though it affects the polluting firm's behavior it does not affect the planner's problem. In this respect the planner's treatment of the intrinsic motivation is

[16] This model has just one polluting firm so does not allow for different degrees of crowding out across different polluters; for example if some firms or if some countries' firms have differing levels of intrinsic motivation. A model extension to include multiple polluters could accommodate this issue.

similar to the "preference purification" approach to behavioral welfare economics.[17] The solution to the planner's problem is thus identical to that in the baseline case. But in the decentralized model, the polluting firm's response to a tax and thus the resulting equilibrium will differ.

The first-order condition for the choice of emissions E in the firm's maximization problem under this specification is $p_x f_E(l_x, E) - \tau - h_E(E, \tau) = 0$. The final term represents how the intrinsic motivation modifies the firm's decision over how much to emit. The planner chooses a pollution tax that accommodates this behavior, given that the planner does not include the intrinsic motivation term in its objective function (the social welfare function). The planner can induce the socially-optimal level of pollution in a decentralized model by setting the following Pigouvian tax rate:

$$\tau = -p_x \sum_{i=1}^{N} \frac{\partial U_i / \partial E}{\partial U_i / \partial x_i} - h_E(E, \tau)$$

The first half of this expression is equal to the Pigouvian tax in the baseline model without intrinsic motivation; the second half represents the planner's response to the fact that the firm exhibits intrinsic motivation. Since $\frac{\partial h}{\partial E} > 0$, this second term decreases the Pigouvian tax rate. Because of intrinsic motivation, the polluting firm itself is at least partially doing the job that the Pigouvian tax should be doing, so the optimal Pigouvian tax need not be as large as it is without intrinsic motivation. The larger the magnitude of the crowd out $\frac{\partial h}{\partial E}$, the smaller the tax needs to be. It is even possible that the second term is larger in magnitude

[17] See Bernheim and Rangel (2009) for a standard methodology adopting this approach, and Infante et al. (2016) for a critique.

than the first term, so that the optimal Pigouvian tax is negative: a pollution subsidy. This would be the case if the intrinsic motivation is so large that the polluting firm "overshoots" and reduces pollution below the socially optimal level even without any pollution tax. Thus a pollution tax can indeed be designed in a way that accommodates the commodification objection, if the commodification objection is fundamentally about the crowding out of intrinsic motivation that commodification entails.

I simulate policy decisions under the intrinsic motivation specification by imposing the following functional form: $h(E, \tau) = mE^n - p\tau^q$. Given this separability between emissions and the emissions tax in the intrinsic motivation cost function, the second part of the function $p\tau^q$ does not affect any outcomes, since the firm treats the pollution tax τ as exogenous. Instead, crowding out comes from the first part of the function, mE^n, where a higher level of emissions chosen by the firm increases this cost. As baseline parameter values I choose $n = 0.5$, which yields interior solutions where the intrinsic motivation creates partial but not total crowd out of the effects of the pollution tax. I vary the value of m from 0 to 0.1, which varies the degree of the intrinsic motivation among the firm and thus the degree of crowding out of the pollution tax.

Figure 9 presents the optimal emissions tax under the planner's problem as I vary the parameter m; this is the expression for the Pigouvian tax presented above. When $m = 0$, there is no intrinsic motivation or crowd out, so the Pigouvian tax is still \$125 per ton. As expected, as the intrinsic motivation is larger, the emissions tax that is needed to maximize social welfare is lower, because the intrinsic motivation of the firm is doing some of the work that the tax would otherwise be doing. There is no clear interpretation of the

magnitude of m, other than to observe that when $m = 0.05$, for example, the intrinsic motivation is such that crowd out causes the tax rate to decrease by about 35% (from $125 per ton to about $80 per ton). The optimal level of emissions that arises from the planner's problem is invariant to the value of m; the intrinsic motivation causes the Pigouvian tax that induces the first-best pollution level to change, but it does not affect the first-best pollution level itself.

Figure 10 presents the emissions tax solving the regulator's problem as I vary m, and Figure 11 presents the resulting emissions level. The regulator's tax displays a qualitatively similar pattern in response to intrinsic motivation as does the planner's tax. There are slight differences in magnitudes, however; as already seen, without intrinsic motivation ($m = 0$) the regulator's tax is slightly lower than the planner's Pigouvian tax. But, at a high degree of intrinsic motivation the regulator's tax is slightly higher than the planner's tax (e.g. when $m = 0.1$, the planner's tax is about $20 per ton, and the regulator's tax is about $30 per ton). This can be explained by Figure 11, which shows that the emissions level from the regulator's tax is not invariant to the degree of crowd out, even though (as mentioned earlier) the first-best emissions level is invariant to it. The regulator who is constrained to use only a pollution tax considers the crowd out response of the firm to the tax, and when doing so allows the emissions level to be slightly lower with higher crowding out. While the magnitude of the change in the emissions level in response to these changes in m is relatively small (varying from about 65% to about 62% of business-as-usual emissions), it is enough to yield a modest change in the regulator's tax relative to the Pigouvian level. As

has been seen throughout all these simulations, the pre-existing income inequality in the economy interacts with all these effects to impact the magnitude of the regulator's tax.

V. Conclusion

I have explored how ethical considerations can be incorporated into an environmental tax. Ethicists have objected to pollution pricing on the grounds of equity, incommensurability, civic responsibility, and commodification. For each of those objections, I have proposed a tax that could be modified from the standard Pigouvian specification (simply equal to marginal external damages) in such a way that accommodates the objection. I perform a series of calibrated numerical simulations to explore how these accommodations affect the magnitude of the optimal tax and the resulting emissions levels. In some cases the accommodation has only a small quantitative effect on the outcomes – for example, equity concerns cause the tax rate to increase by a modest amount. In other cases the accommodation can have a large effect – for example, under a reasonable magnitude of a civic responsibility accommodation, the resulting tax rate can be about five times larger than the Pigouvian level.

This study is meant to be a starting point for incorporating these ethical concerns into the design of real-world pollution pricing policies like carbon taxes. Much more work can be done to explore these issues further. In the theoretical results presented here, many simplifying assumptions were made to yield interpretable results, and these assumptions could be modified in future work. For example, the objection from incommensurability was accommodated with a simple constraint on the total level of

emissions; many alternative ways of accommodating this objection could be found. Likewise, only one market failure (a global pollution externality) is modeled; extensions could consider co-pollutants or other market failures. The quantitative simulation results arise from a relatively simple computational model calibrated to the U.S. economy, with just five representative consumers and two representative firms. More detailed structural modeling of the economy, as in a multi-sector computable general equilibrium model, could yield more informative quantitative results. Furthermore, the specification of environmental damages is simple (linear), and the model is static; a dynamic model with a more realistic environmental component, like an integrated assessment model, could provide additional insights. I have considered each ethical concern independently of the others; it is possible that more insights can be gleaned by addressing multiple ethical concerns simultaneously.

Nevertheless, this study is the first economic model to seriously engage with this literature from environmental ethics and the philosophy of environmental economics in this way. Clearly the policy implications are directly related to the design of environmental pricing policies like carbon taxes. Standard economic models, including computable general equilibrium and integrated assessment models, generally ignore these ethical concerns, though policymakers and the populace may care about them deeply. The results here will hopefully inform the design of these policies in a way that can solve environmental problems in an economically efficient way while not raising ethical objections.

References

Adler, Matthew D. "Benefit–cost analysis and distributional weights: an overview." *Review of Environmental Economics and Policy* (2016).

Aldred, Jonathan. 2012. “The Ethics of Emissions Trading.” *New Political Economy* 17 (3): 339–60.

Anderson, Elizabeth. 1995. *Value in Ethics and Economics*. Harvard University Press.

Argenziano, Rossella, and Itzhak Gilboa. "Perception-theoretic foundations of weighted utilitarianism." *The Economic Journal* 129, no. 620 (2019): 1511-1528.

Barrage, Lint, and William Nordhaus. "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model." *Proceedings of the National Academy of Sciences* 121, no. 13 (2024): e2312030121.

Bernheim, B. Douglas, and Antonio Rangel. "Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics." *The Quarterly Journal of Economics* 124, no. 1 (2009): 51-104.

Bovenberg, A.L. and De Mooij, R.A., 1994. Environmental levies and distortionary taxation. *The American Economic Review*, *84*(4), pp.1085-1089.

Buchholz, Wolfgang, and Jan Schumacher. "Discounting and welfare analysis over time: Choosing the η." *European Journal of Political Economy* 26, no. 3 (2010): 372-385.

Caney, Simon, and Cameron Hepburn. 2011. “Carbon Trading: Unethical, Unjust and Ineffective?” *Royal Institute of Philosophy Supplement* 69 (October):201–34.

Congressional Budget Office (CBO). 2008. “Historical Effective Tax Rates, 1979 to 2005: Supplement with Additional Data on Sources of Income and High-Income

Households." Available at https://www.cbo.gov/sites/default/files/110th-congress-2007-2008/reports/12-23-effectivetaxratesletter.pdf

Cowell, F. A., and K. Gardiner. "Welfare weights (STICERD)." *Economics Research Paper* 20 (1999): 1999.

Deb, Kalyanmoy. "Multi-objective optimisation using evolutionary algorithms: an introduction." In *Multi-objective evolutionary optimisation for product design and manufacturing*, pp. 3-34. London: Springer London, 2011.

Del Campo, Stellio, David Anthoff, and Ulrike Kornek. "Inequality aversion for climate policy." *Review of Environmental Economics and Policy* 18, no. 1 (2024): 96-115.

Fehr, Ernst, and Gary Charness. 2025. "Social Preferences: Fundamental Characteristics and Economic Consequences." *Journal of Economic Literature* 63 (2): 440–514**.**

Fowlie, Meredith, 2010. Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review*, *100*(3), pp.837-869.

Frey, B. and Stutzer, A. (2008), "Environmental Morale and Motivation," in A. Lewis (ed)., *The Cambridge Handbook of Psychology and Economic Behaviour*, pp. 406–29 (Cambridge: CUP).

Goodin, Robert E. 1994. “Selling Environmental Indulgences.” *Kyklos* 47 (4): 573–96.

Heutel, Garth, and J. Paul Kelleher. 2024. “The Ethics of Pollution Pricing.” National Bureau of Economic Research working paper 33300.

Howarth, Richard B. "Climate rights and economic modeling." In *The Long-Term Economics of Climate Change: Beyond a Doubling of Greenhouse Gas Concentrations*, pp. 315-336. Emerald Group Publishing Limited, 2001.

Infante, Gerardo, Guilhem Lecouteux, and Robert Sugden. "Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics." *Journal of Economic Methodology* 23, no. 1 (2016): 1-25.

Jacobs, Bas, and Ruud A. De Mooij. "Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax." *Journal of Environmental Economics and Management* 71 (2015): 90-108.

Jaffe, Adam B., Richard G. Newell, and Robert N. Stavins. "A tale of two market failures: Technology and environmental policy." *Ecological economics* 54, no. 2-3 (2005): 164-174.

Janssen, Ron. *Multiobjective decision support for environmental management*. Vol. 2. Springer Science & Business Media, 2012.

Kelleher, J. Paul. *The Social Cost of Carbon: Ethics and the Limits of Climate Change Economics*. Oxford University Press, 2025.

Llavador, Humberto, John E. Roemer, and Joaquim Silvestre. "A dynamic analysis of human welfare in a warming planet." *Journal of Public Economics* 95, no. 11-12 (2011): 1607-1620.

Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R. Green. *Microeconomic Theory*. New York: Oxford University Press, 1995.

Metcalf, G.E., 2003. Environmental levies and distortionary taxation: Pigou, taxation and pollution. *Journal of Public Economics*, *87*(2), pp.313-322.

Mintz-Woo, Kian. 2024. “Carbon Tax Ethics.” *WIREs Climate Change* 15 (1): e858.

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

Page, Edward A. 2012. “The Hidden Costs of Carbon Commodification: Emissions Trading, Political Legitimacy and Procedural Justice.” *Democratization* 19 (5): 932–50.

Parry, I.W., 1995. Optimal pollution taxes and endogenous technological progress. *Resource and energy economics*, *17*(1), pp.69-85.

Paul, Anthony C., Blair Beasley, and Karen Palmer. 2013. “Taxing electricity sector carbon emissions at social cost.” Resources for the Future Discussion Paper 13-23-REV.

Phaneuf, Daniel J., and Till Requate. *A course in environmental economics: theory, policy, and practice*. Cambridge University Press, 2016.

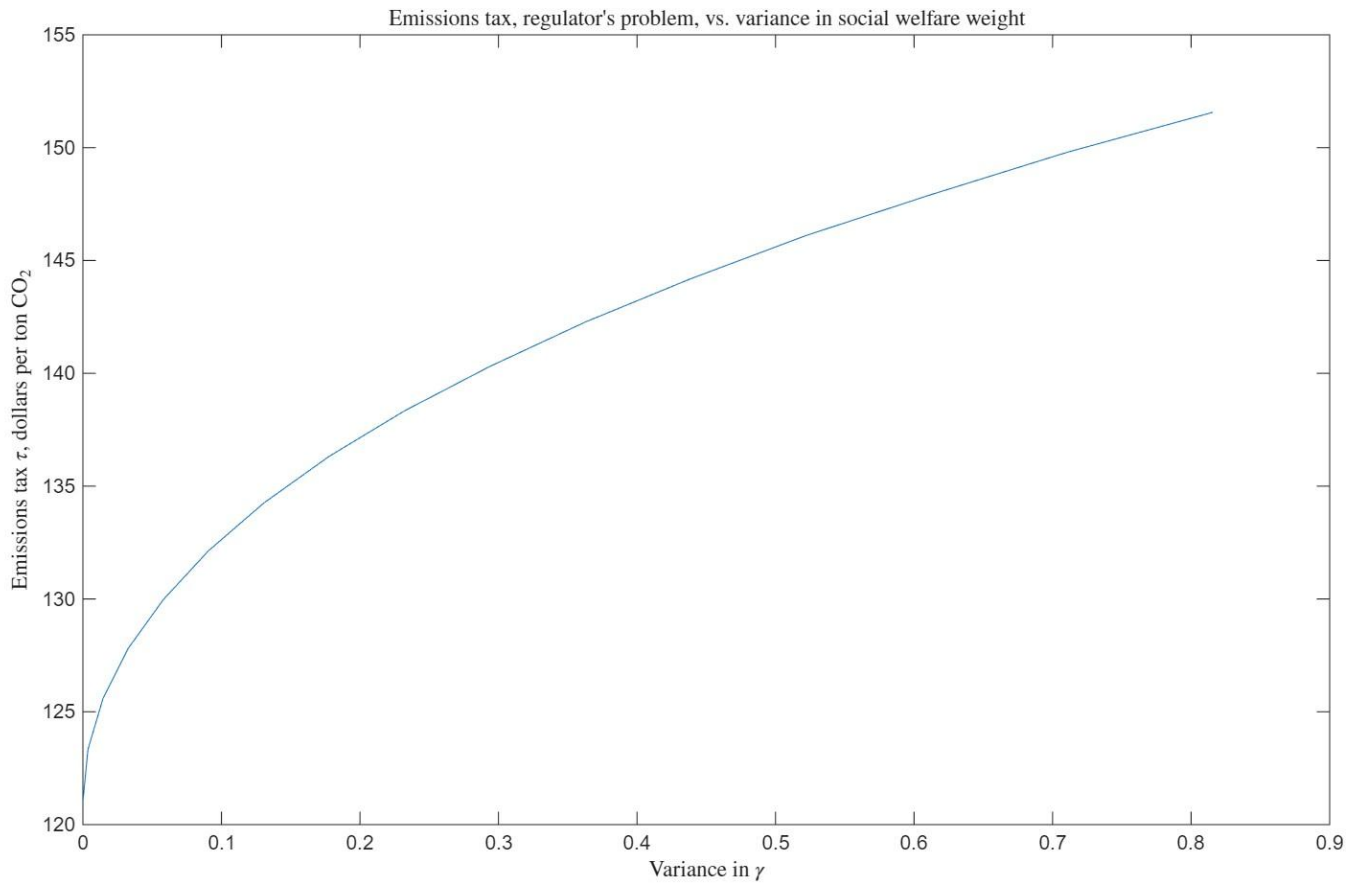
Sandel, Michael J. 2012. *What Money Can’t Buy: The Moral Limits of Markets*. Macmillan.

Satz, Debra. 2010. *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*. Oxford University Press.

Smart, John Jamieson Carswell, and Bernard Williams. *Utilitarianism: For and against*. Cambridge University Press, 1973.

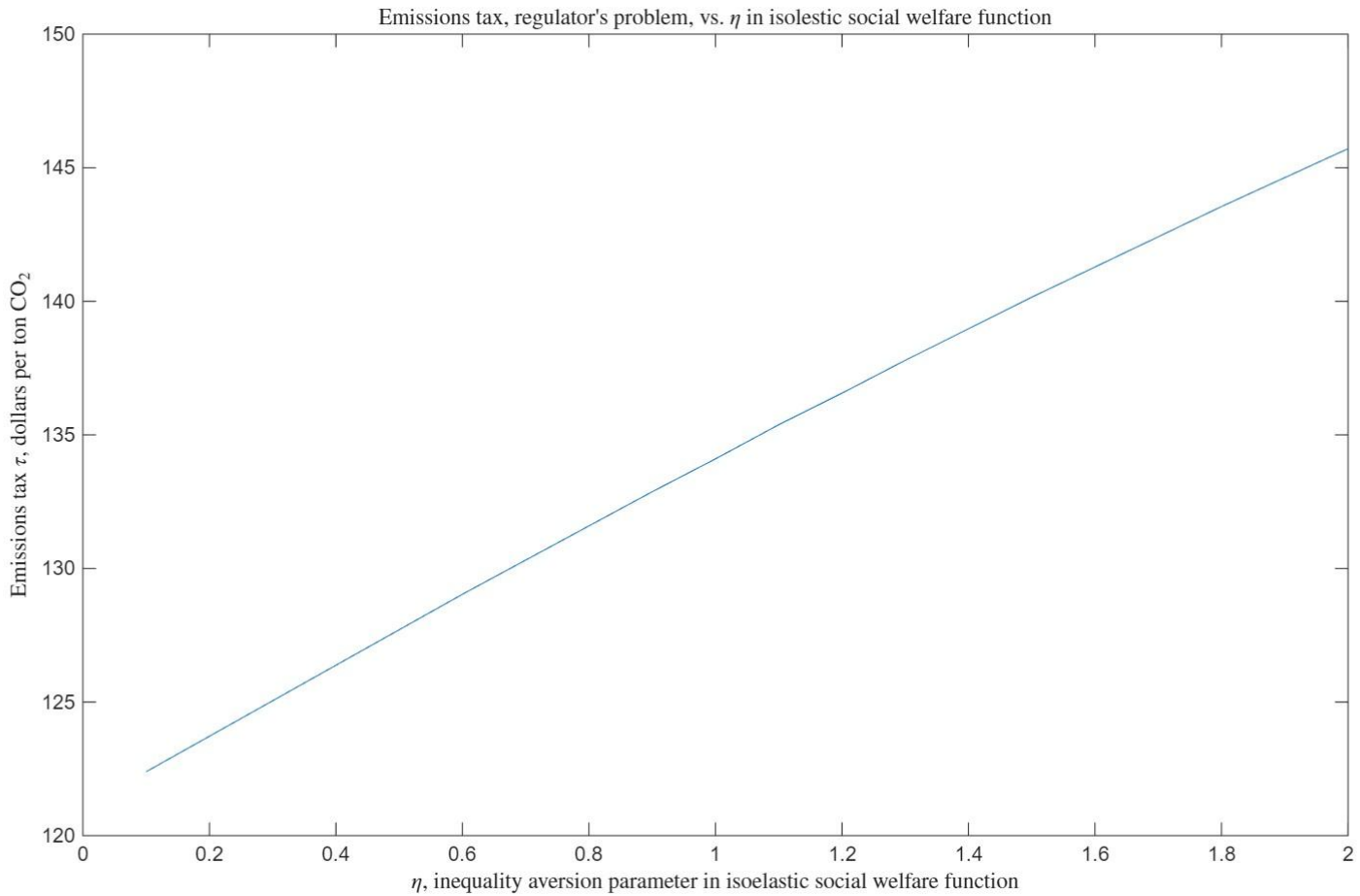
West, Sarah E., and Roberton C. Williams III. "Estimates from a consumer demand system: implications for the incidence of environmental taxes." *Journal of Environmental Economics and Management* 47, no. 3 (2004): 535-558.

Figure 1



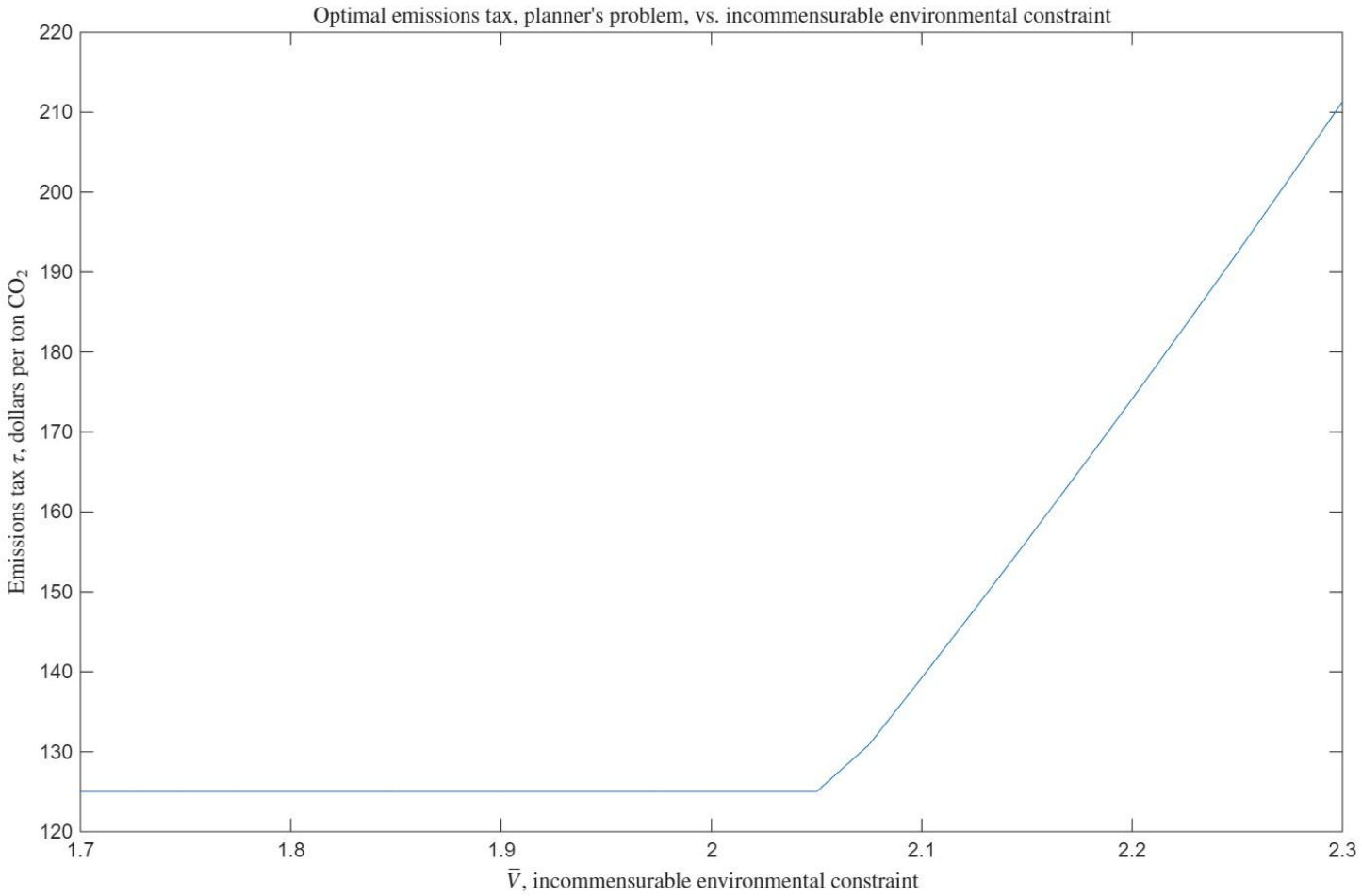
Notes: The x-axis is the variance in the vector of social welfare weights γ, ranging from zero (where all the γ_i values equal one) to 0.82 (where the γ_i for the richest quintile is 0.25 and the γ_i for the poorest quintile is 2.5). The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Figure 2



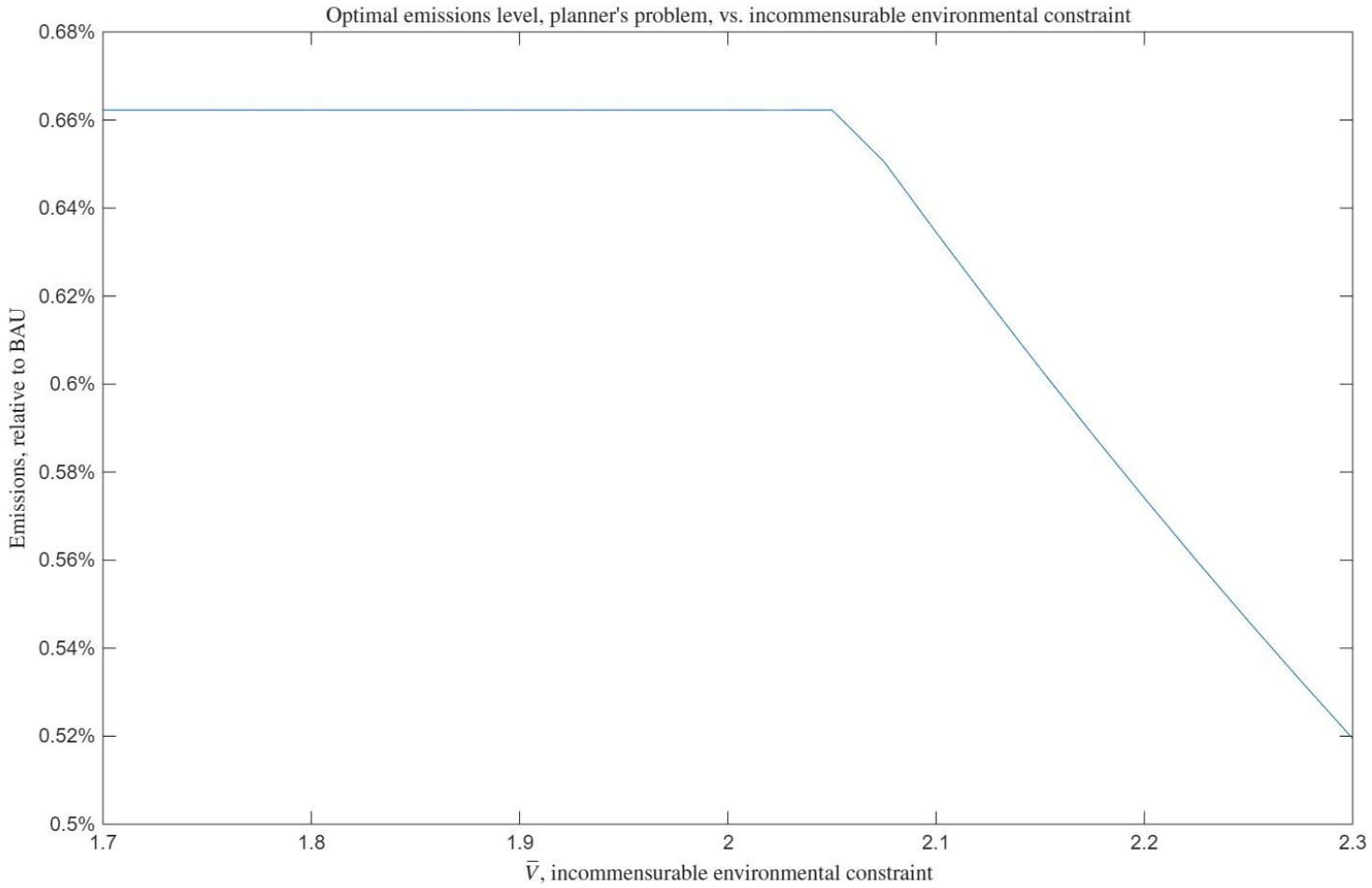
Notes: The x-axis is η, the inequality aversion parameter in the isoelastic social welfare function. The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Figure 3



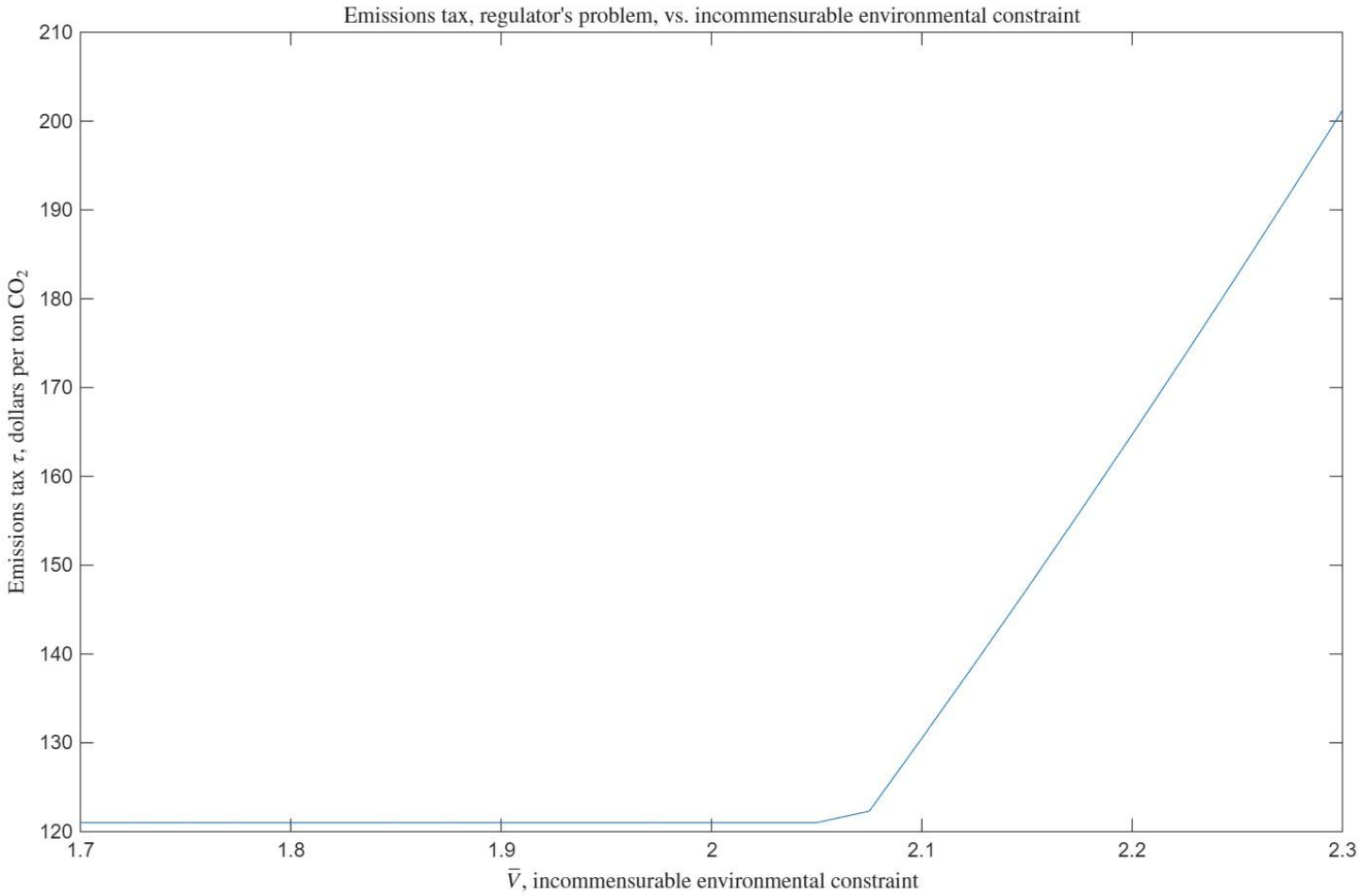
Notes: The x-axis is $\bar{V}$, the value of the minimum environmental quality constraint in the incommensurability specification. The y-axis is the emissions tax rate that solves the planner's problem, i.e., the optimal emissions tax, in 2019 dollars per ton of carbon dioxide.

Figure 4



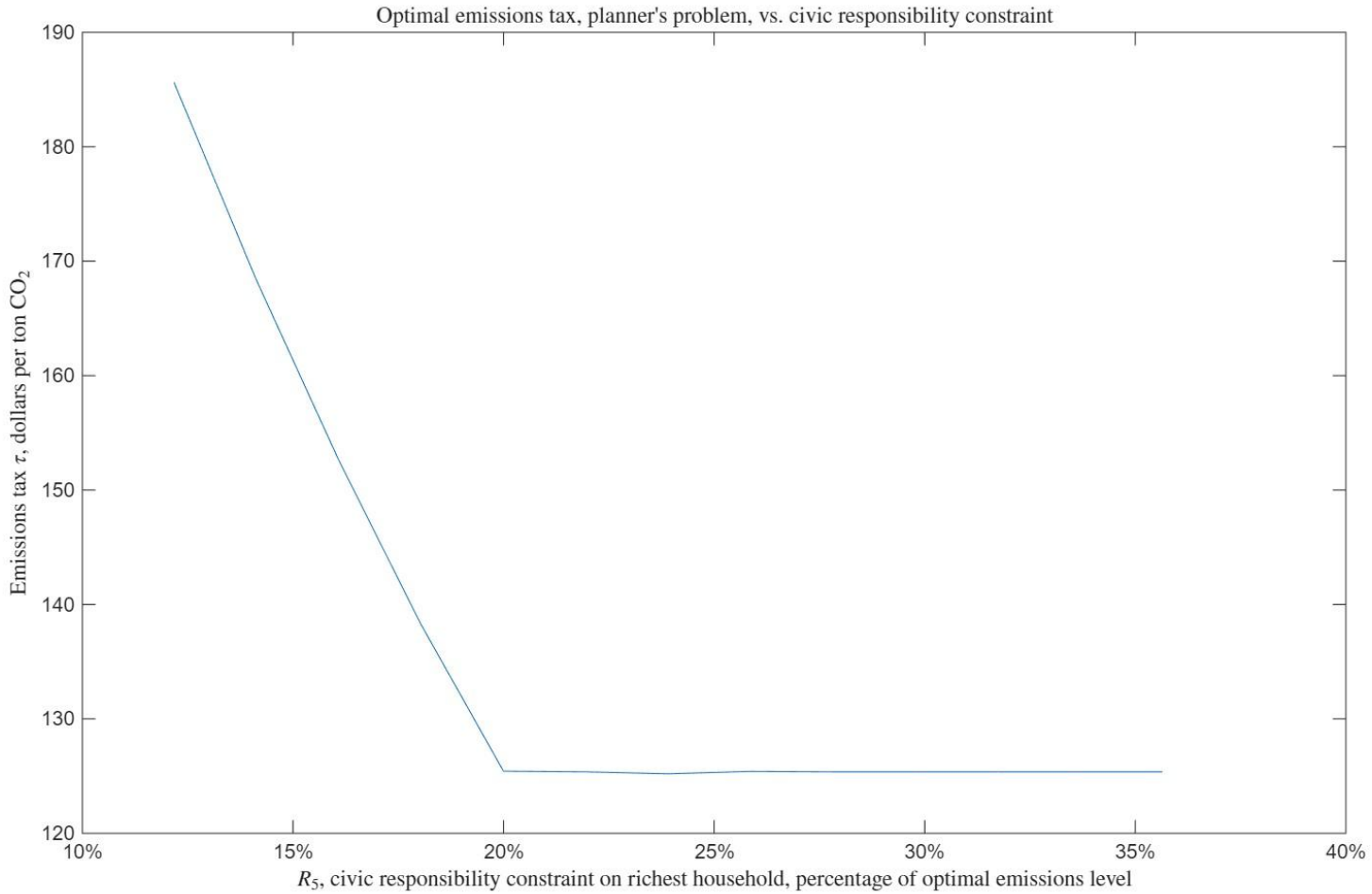
Notes: The x-axis is $\bar{V}$, the value of the minimum environmental quality constraint in the incommensurability specification. The y-axis is the optimal quantity of emissions from the planner's problem, expressed as a percentage of the business-as-usual emissions level.

Figure 5



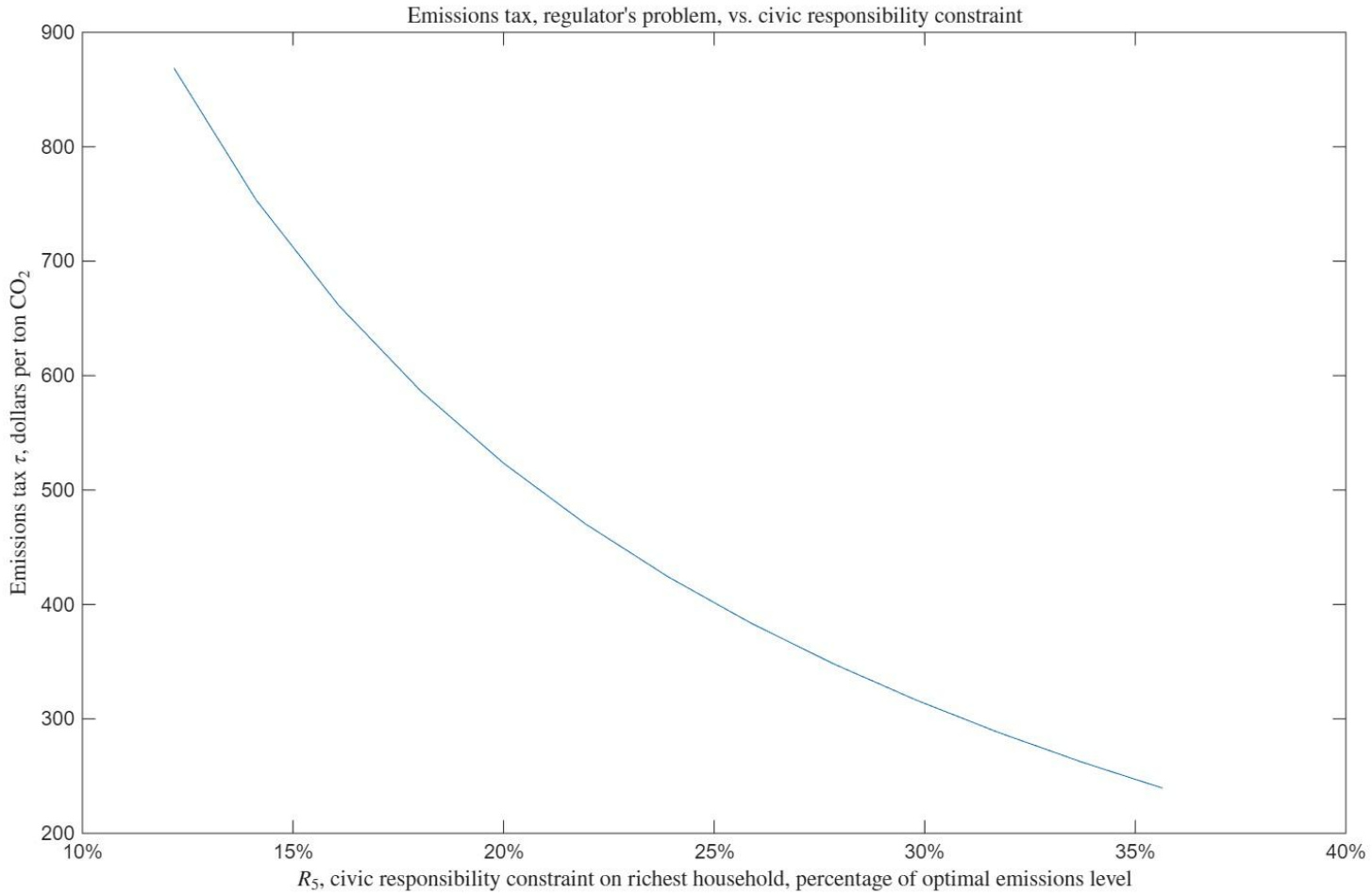
Notes: The x-axis is $\bar{V}$, the value of the minimum environmental quality constraint in the incommensurability specification. The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Figure 6



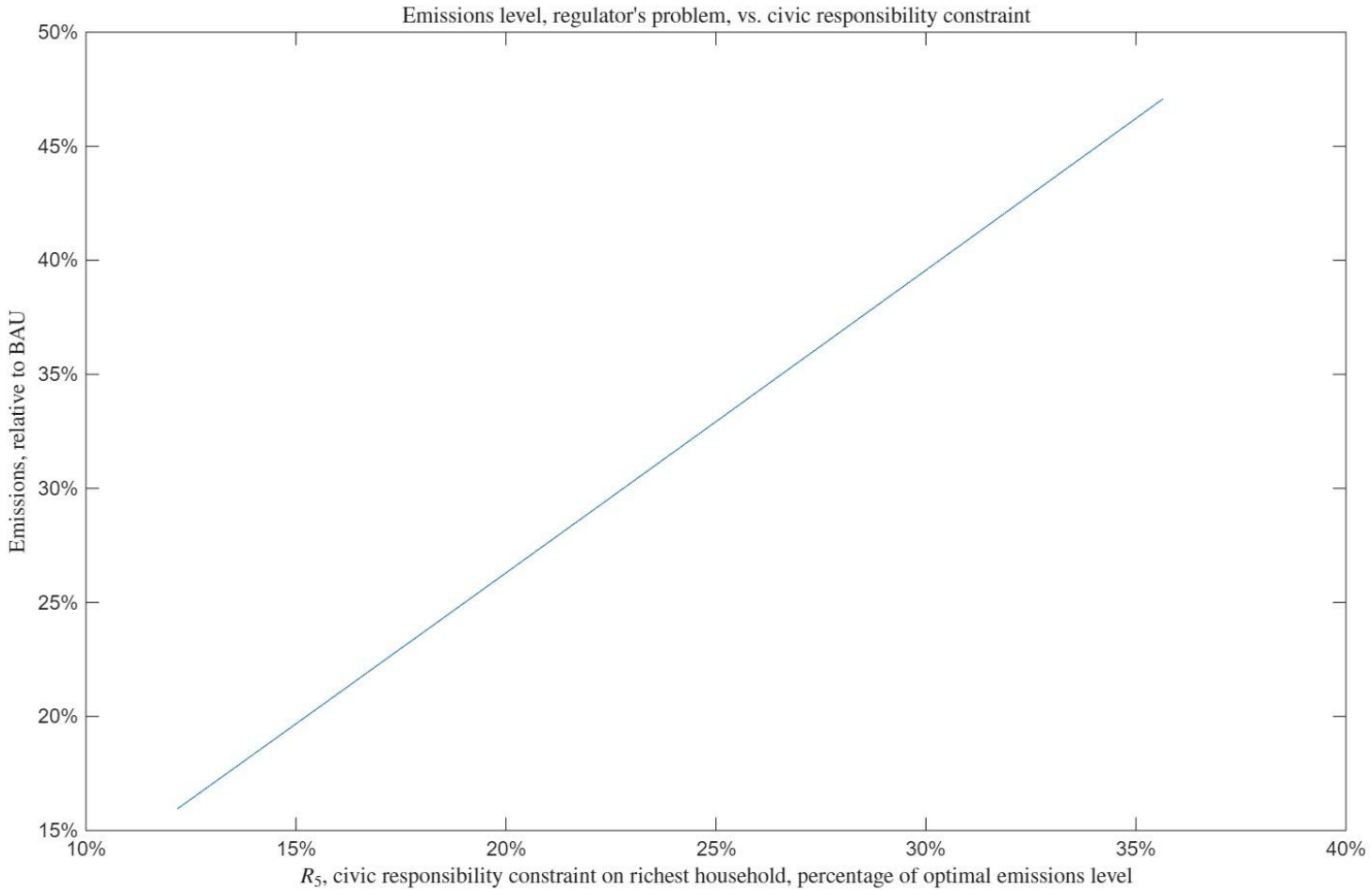
Notes: The x-axis is R_5, the value of the civic responsibility on the richest household, expressed as a percentage of the overall optimal emissions level (solving the planner's problem). The constraints on the other four households are kept constant at 20% of the overall optimal emissions level. The y-axis is the emissions tax rate that solves the planner's problem, i.e., the optimal emissions tax, in 2019 dollars per ton of carbon dioxide.

Figure 7



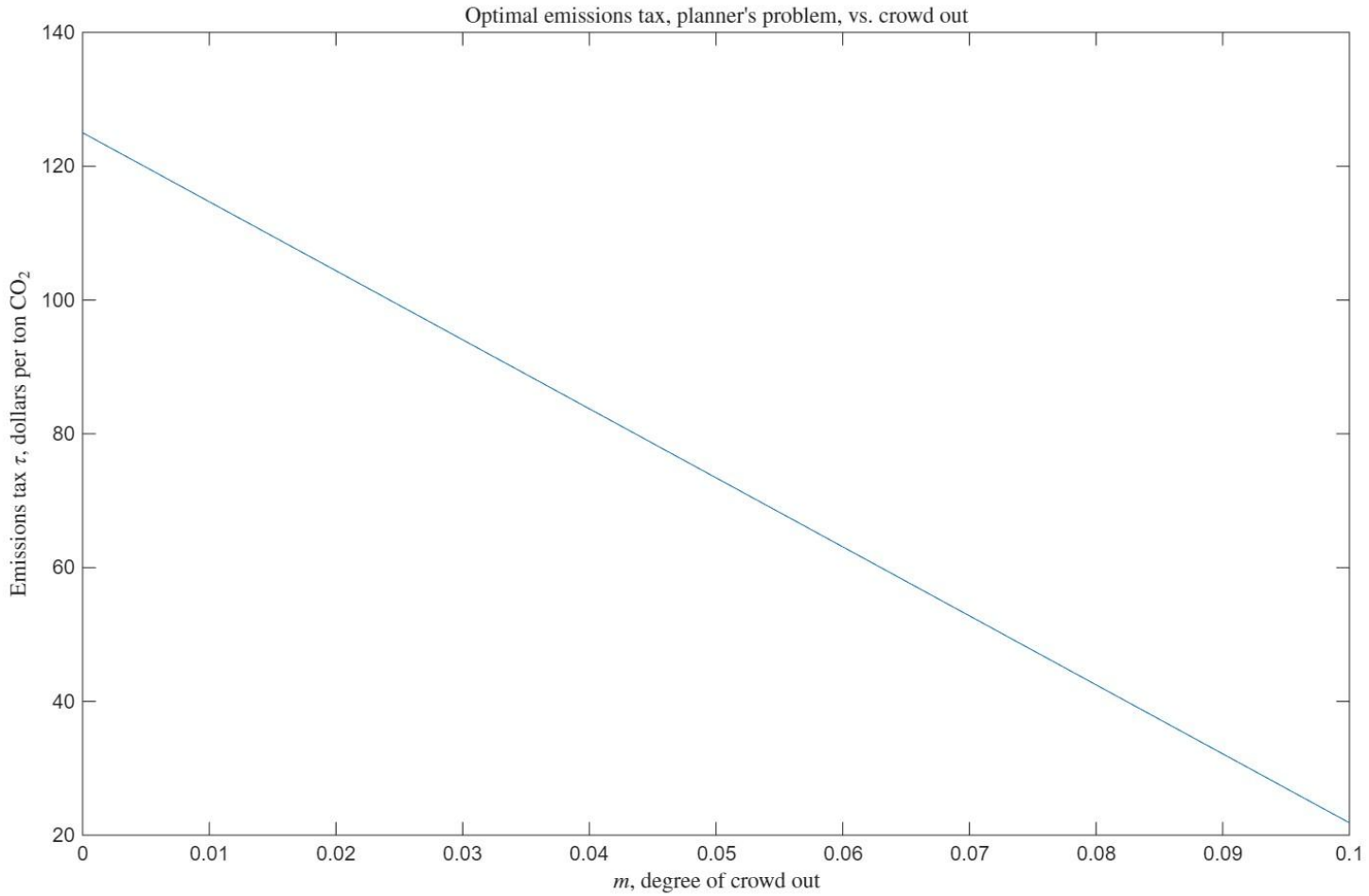
Notes: The x-axis is R_5, the value of the civic responsibility on the richest household, expressed as a percentage of the overall optimal emissions level (solving the planner's problem). The constraints on the other four households are kept constant at 20% of the overall optimal emissions level. The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Figure 8



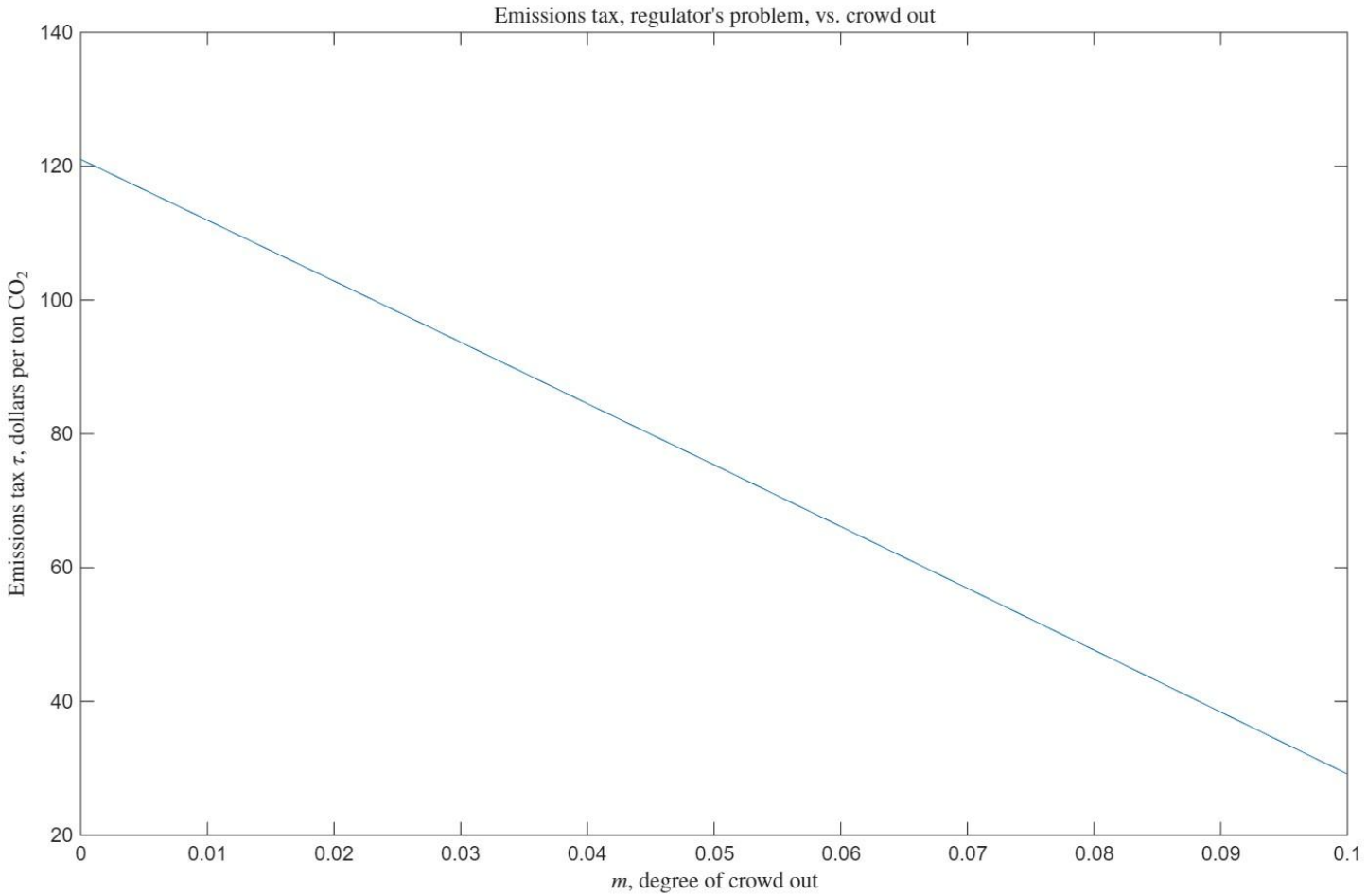
Notes: The x-axis is R_5, the value of the civic responsibility on the richest household, expressed as a percentage of the overall optimal emissions level (solving the planner's problem). The constraints on the other four households are kept constant at 20% of the overall optimal emissions level. The y-axis is the quantity of emissions that results from the solution to the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Figure 9



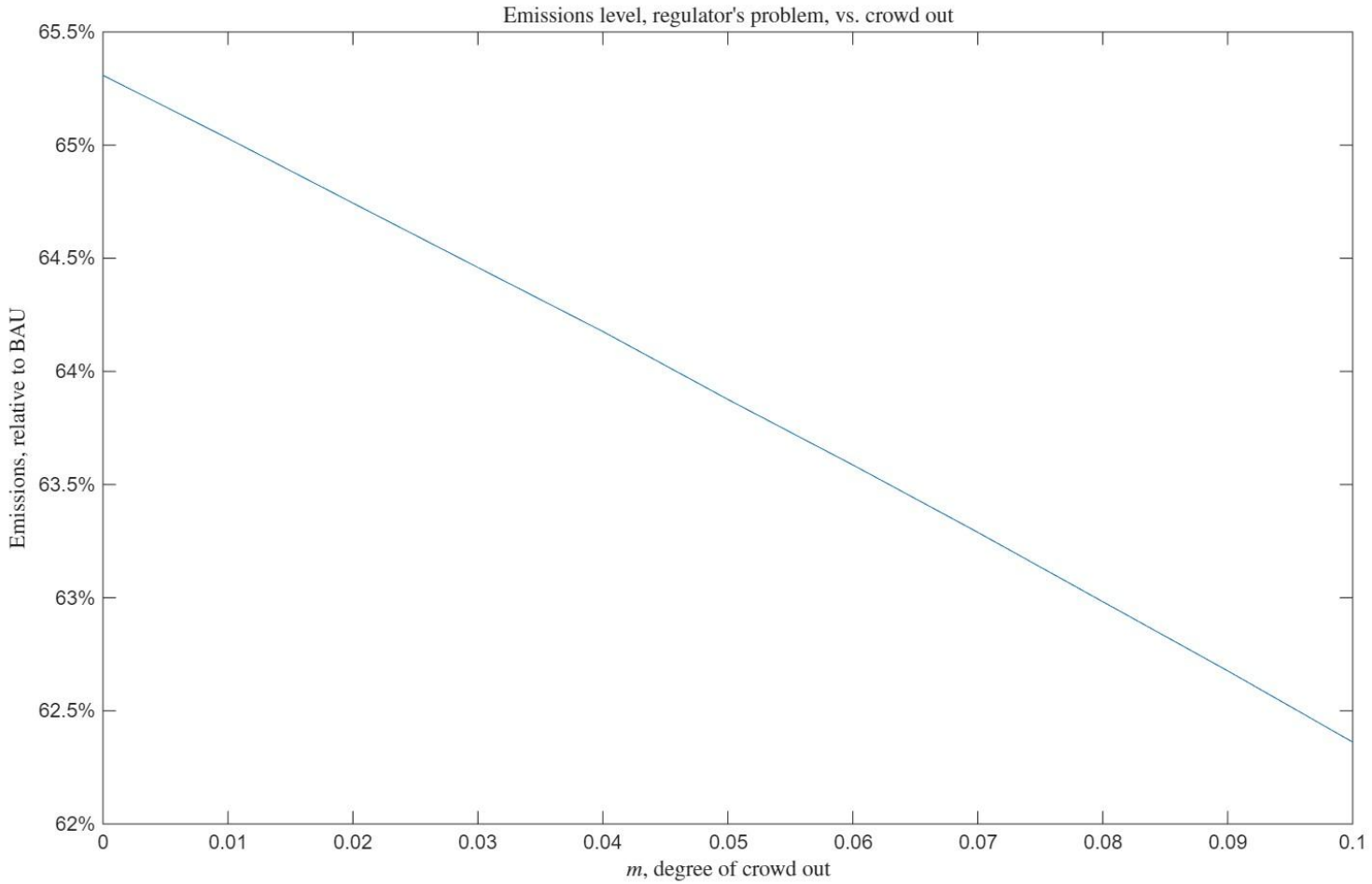
Notes: The x-axis is m, a parameter governing the degree of intrinsic motivation causing crowding out in the firm's objective function. The y-axis is the emissions tax rate that solves the planner's problem, i.e., the optimal emissions tax, in 2019 dollars per ton of carbon dioxide.

Figure 10



Notes: The x-axis is m, a parameter governing the degree of intrinsic motivation causing crowding out in the firm's objective function. The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Figure 11



Notes: The x-axis is m, a parameter governing the degree of intrinsic motivation causing crowding out in the firm's objective function. The y-axis is the quantity of emissions that results from the solution to the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Appendix

A.I. Additional Background

A.I.a. Equity

It has long been argued that pollution pricing may increase efficiency at the cost of equity, since poorer people tend to spend a higher fraction of their income on pollution-intensive goods like electricity and gasoline. For example, see Glaeser et al. (2023) on the regressivity of gasoline taxes. But see Bruegge et al. (2019), who find that command-and-control regulations can also be regressive. The design of the revenue return of the pollution tax has important distributional effects; see West and Williams (2004).[18]

Many theoretical and computational studies have evaluated environmental tax policy in the context of distributional concerns. A recent review of the economics literature on inequality and the environment is Drupp et al. (2025). These studies date back at least to Sandmo (1975), who considers optimal pollution taxation where the social welfare function is unweighted utilitarian and people differ in their productivity. His theoretical model demonstrates how the optimal pollution tax is affected by this heterogeneity (see the first term of his equation 35), and he demonstrates that a higher optimal pollution tax is more likely when high-income people are polluters and low-income people are "pollutees" (those harmed by pollution). This theoretical exploration continues in Cremer et al. (1998), who also consider an unweighted utilitarian social welfare function and consumers

[18] See also Baranzini et al. (2017) or Klenert et al. (2018) for discussions of political economy considerations related to carbon pricing, a dimension which I do not address in this paper.

differing in productivity, and who also allow consumers to differ in tastes. Their focus is a justification for a non-linear tax on the externality. Neither of these theoretical papers explicitly models social preferences for equality through for example a weighted utilitarian social welfare function.

Other studies have extended theoretical results to numerical simulations to gain further intuition. Cremer et al. (2003) extend their earlier theoretical paper by calibrating and simulating their model for the French economy, considering four consumer types. In this paper, the social welfare function is isoelastic, allowing an inequality aversion parameter to be included. They quantify how the welfare gains of environmental taxes are distributed across the four consumer groups under different scenarios and assumptions. Generally, they find that redistributive concerns cause the optimal pollution tax to be much below marginal external damages. Abrell et al. (2018) also combine a theoretical model with computational simulations to explore pollution taxes under distributional concerns; in their paper the focus is on how taxes can be differentiated across different sectors to maximize a social welfare function that cares about equity. Jacobs and van der Ploeg (2019) consider consumers heterogeneous in productivity and pollution damages, and they find that the optimal pollution tax is higher if pollution damages disproportionately hurt the poor and lower if the poor spend disproportionately more on polluting goods. Van der Ploeg et al. (2025) consider how the choice of revenue return for carbon taxes can influence both efficiency and equity. While many of these studies consider complications that I ignore, like heterogeneity in consumer productivity, the purpose of my model is to

explore ethical issues beyond just equity concerns, so I use a very simple specification on this dimension.

Other studies in public and environmental economics have considered tax design under equity concerns motivated by the theory of justice as fairness introduced in Rawls (1971). The incorporation of a Rawlsian framework for optimal tax policy in the economics literature extends back at least to Howe and Roemer (1981), and the early literature is summarized in Walraevens (2023). Fleurbaey and Maniquet (2018) argue that many concepts of policy desiderata, including Rawlsian concepts of justice, can be incorporated into normative economic analysis through an appropriate social welfare function. For example, the iso-elastic welfare function is often used and called a "Rawlsian" social welfare function; as the inequality aversion parameter goes to infinity, it becomes a maximin function. However, Walraevens (2023) argues that such a welfarist framework is in fact unfaithful to Rawls. In the context of environmental and climate policy, Becker (1982) theoretically models optimal environmental policy in continuous time according to a Rawlsian maximin criterion, and d'Autume et al. (2010) consider a model similar to an integrated assessment model but with a maximin social welfare function. DeCanio et al. (2022) use the minimax-regret criterion in climate policy analysis, a criterion that shares much with Rawlsian notions of equity. In section III.a. I extend my model to allow for Rawlsian-style social welfare functions.

A.I.b. Incommensurability

The argument against pollution pricing based on incommensurability is forcefully expressed by Elizabeth Anderson (1995). She argues that environmental goods (and other types of goods) should be valued "intrinsically," as opposed to only instrumentally, and further that this type of valuation renders the commensurability between environmental and market goods morally unjustified. The resulting incommensurability casts moral doubt on any justification for pollution pricing (or any other environmental policy) based on a cost-benefit analysis that directly compares environmental and market goods. Such commensurability is underlying normative integrated assessment models (IAMs), like the DICE model. The imposition of the commensurability between environmental and market goods extends at least as far back as Krutilla (1967); see Raz (1985) for an early summary of arguments.

This argument for incommensurability reflects *value pluralism* – the notion that different things can have different, possibly incommensurable, types of values – as opposed to *value monism*, which is implicitly assumed in much normative environmental economic analysis. The field of ecological economics is generally more amenable to value pluralism (Gowdy and Erickson 2005, O'Neill 2017). Martinez-Alier et al. (1998) argue that incommensurability of values is a characteristic feature of ecological economics; Thoma (2023) has argued for expanding value pluralism across the social sciences. As even the defenders of value pluralism admit, the way in which these different types of values can be evaluated and used to make policy decisions, while still respecting their incommensurability, is unclear (Anderson 2014).

Incommensurability is clearly an issue for a Pigouvian justification of pollution taxation, which explicitly relies on comparing environmental and market goods' values. But incommensurability may not necessarily yield an objection to any use of pollution taxation, if the pollution tax or its defender makes no claim about the tax being "optimal" or Pigouvian. That is, suppose that a pollution reduction goal has been decided, through a process that respects the incommensurability of different types of goods. Then, regulators choose to meet that pollution reduction goal through a pollution tax, to take advantage of the cost-effectiveness benefits that a pollution tax promise. Does this use of a pollution tax still create issues for those worried about incommensurability? According to Aldred (2012), it indeed does. His argument is that putting a price on pollution implicitly imposes commensurability via the implication of social norms. Others disagree; see Page (2011). The tax that I propose overcomes the Pigouvian type of commensurability that the standard neoclassical model imposes, though it still falls victim to Aldred's argument since it does in fact put a price on pollution.

The tax proposed in the text to accommodate incommensurability is reminiscent of the "standard-and-tax" approach developed by some researchers in the mid-20th century and described in Berta (2020). That approach acknowledged the difficulty (near-impossibility) of calculating a value for marginal social damages that would be required in a Pigouvian tax. Instead, they proposed independently imposing "acceptability standards" for pollution levels and then using a pollution tax to reach that level. Defenders of that approach were unlikely relying on philosophical arguments against commensurability in

theory and instead were most concerned with the difficulty of making comparisons in practice. Still, this approach bears similarities with the approach taken here.

One might still object that the expression representing this tax in fact violates incommensurability arguments, because the value of the environment is directly compared with values over market goods in that expression. Fair enough. But it must be remembered that the expression arises from a social welfare function that does not compare those two types of values. To yield an optimal tax expression, this model relied on a very specific and mathematically-convenient form of incommensurability: the planner maximizes welfare over consumption goods conditional on environmental quality meeting a certain minimum. Other ways of modeling incommensurability may not yield such an interpretable form of the optimal pollution tax. But this purpose of this exercise was to demonstrate that, in principle, it is possible to design a pollution tax that respects the philosophical position of the incommensurability between environmental and other types of goods. Intuitively, the pollution tax respects that position because it is derived based on a social welfare function respecting that position; the tax is merely the instrument used to achieve the socially-optimal outcome. No doubt some philosophers would disagree with this interpretation. As discussed earlier, Aldred (2012) argues that the enactment of the pollution tax itself implicitly imposes commensurability.[19]

A.I.c. Collective Responsibility

[19] Kelleher (2025) discussed this issue and many other ethical issues related to the development of a social cost of carbon.

Goodin (1994) compares pollution pricing to the Catholic Church's history of selling *indulgences*: allowing sinners to escape the moral repercussions of their sins by paying the church. If polluting, or at least polluting above some acceptable level, is the moral equivalent of a sin, then allowing a higher level of pollution in exchange for a payment to the government is morally analogous to religious indulgences (and objectionable).

Mintz-Woo (2022) discusses this objection and refers to it as the *solidarity argument*. He in particular applies the objection to emissions trading between rich and poor countries; if rich countries are not reducing their emissions but instead are paying poor countries to do so through buying emissions permits, "not everyone will be contributing in an appropriate manner." (p. 6) The argument, though, is not specifically about rich countries vs. poor countries, nor about rich polluters vs. poor polluters within a country, but rather more generally about obligations that (perhaps) we all have and how the structure of emissions pricing allows some of us to evade those obligations.

Collective responsibility as it specifically relates to climate change has been explored by Vanderheiden (2011) and Hormio (2023), though neither study addresses the relationship between collective responsibility and carbon pricing. Levi (2021) conducts a survey on public support for carbon taxes and finds that the most important determinant of support is the sense of personal responsibility for reducing climate change. Similarly, Obradovich and Guenther (2016) find from experimental evidence that emphasizing collective responsibility for the causes of climate change increases pro-climate donations and intentions to reduce emissions. However, I know of no study that attempts to

incorporate the notion of collective responsibility into the design of pollution pricing, as I do.

A.I.d. Commodification

Critiques of pollution pricing via the commodification argument are found in Paton and Bryant (2023) and Berta (2023). A review of the commodification of nature literature, focusing primarily within ecological economics, is in Smessaert et al. (2020).

There is a debate over whether in fact pollution pricing *does* commodify the environment. Caney and Hepburn (2011) provide an argument that it does not (necessarily) do so, since well-designed emissions trading systems grant merely *use rights* to the natural environment rather than *property rights*. Aldred (2012) rejects this argument and this distinction; a pollution pricing policy is objectionable because it commodifies the environment via the "rival account of value" (p. 346) that it endorses. Some empirical evidence on the matter is provided by Kim (2024), who conducts an experiment involving a fictitious pollutant and finds that hypothetical market-based regulation of the pollutant does not affect subjects' moral assessment of the pollutant.

If the commodification argument holds, then a policy that puts a price on pollution through a tax or a tradable permit system is morally objectionable regardless of the efficiency gains. This moral objection to pollution pricing is more "essentialist" than the others, in the sense that it is the pricing itself that is objectionable, rather than some outcome or implication of the pricing. This argument is similar to another ethical objection, identified by both Caney and Hepburn (2011) and Aldred (2012), which is the

"pricing what should not be priced" argument. For the purposes of this paper and the theoretical models, I treat the commodification argument and the "pricing what should not be priced" argument as identical to each other, though see the two earlier-cited papers for a discussion on their differences. For my purposes, both arguments can take the "essentialist" position that it is simply the commodification itself, or the pricing itself, that is morally objectionable regardless of consequences, e.g. efficiency or equity impacts.

But another aspect of the anti-commodification argument is not essentialist. It relates to the impact that commodification has on crowding out intrinsic motivations for protecting the environment. While Heutel and Kelleher (2024) treat commodification and crowding out as distinct objections to emissions pricing, for my purposes here I consider crowding out as a potential mechanism through which the commodification argument can be made. That is, one can claim that pollution pricing immorally commodifies nature based on the fact that this commodification of nature crowds out individuals' intrinsic motivations to protect nature; this argument can be seen in Frey and Stutzer (2008). As I explore, this non-essentialist anti-commodification argument will be able to be accommodated with a pollution tax that recognizes the crowding out of intrinsic motivation and is designed with it in mind. An essentialist commodification argument, however, cannot be accommodated with any pollution price policy.

A.II. Calibration Details

I use the following parameter values in the calibrated simulations (summarized in Appendix Table A1). I use data on income across the U.S. population to calibrate the

income share parameters θ. The report CBO (2008) provides data on total income by households across capital, labor, and transfer income, with households divided by income quintile (and into finer divisions among the highest quintile, which I ignore and aggregate). I use that report's Table 5 and data for 2005 (the most recent year available). I ignore transfer income and taxes, and instead I focus on pretax labor and capital income. For each quintile, I set the fraction of labor income θ_{il} to equal the fraction of total wage income earned by that quintile from CBO (2008). For the fractions of firm profits allocated to each quintile for the dirty firm (θ_{ix}) and clean firm (θ_{iz}), I set these to equal the fraction of total capital income earned by that quintile from CBO (2008). Capital income in CBO (2008) is not disaggregated by source, so I cannot identify the difference in income from clean vs. dirty firms, and so for each quintile i, $\theta_{ix} = \theta_{iz}$. The labor income share varies from $\theta_{1l} = 3\%$ to $\theta_{5l} = 50\%$, and the capital income share varies from $\theta_{1x} = 1\%$ to $\theta_{5x} = 87\%$. For the income from pollution tax revenue $\theta_{i\tau}$, I impose that the revenues are returned equally across the income quintiles, so that $\theta_{i\tau} = \frac{1}{N} = \frac{1}{5}$ for each quintile.[20]

For the production function shares for the dirty good sector, Hassler et al. (2012) report that the factor share of energy in the US economy ranges from about 2% to 6% from 1950 through 2010 (their Figure 1). In my model, if we equate the energy share with the pollution (E) share, this share corresponds to $1 - \kappa$ (the exponent on E in dirty good production). But, because my model includes labor but not capital, I scale up $1 - \kappa$ so that

[20] The model only allows for pollution tax revenues to be returned as lump-sum transfers, rather than as cuts in other taxes like labor or capital taxes (Fried et al. 2024), or to be spent on energy efficiency or other environmental projects as in the EU ETS (Wiese et al. 2020). Furthermore I only consider lump-sum transfers distributed according to the fixed $\theta_{i\tau}$ values, rather than the regulator being allowed to adjust this distribution (Budolfson et al. 2021).

the energy/labor ratio is about equal to that reported from the data. I thus set $\kappa = 0.7$. For symmetry, I use the same exponent on labor in the clean good production sector and set $\mu = 0.7$. The exponents on the clean and dirty good, respectively, in the utility function are $\alpha = 0.3$ and $\omega = 0.4$. These parameter values are arbitrary but are chosen to ensure an interior solution in the planner's problem and the decentralized economy. Total labor supply is normalized so that $l = 1$.

The parameter describing the disutility of pollution in the utility function β is calibrated based on two stylized facts. First, at the efficient outcome, the level of emissions is 30% lower than it is at the business-as-usual (no pollution tax) outcome. The 30% target comes from recent simulations of "optimal" policy under the DICE-2023 model, where in 2050 optimal emissions are about 30% lower than business-as-usual (Barrage and Nordhaus 2024, Figure 1). Second, Paul et al. (2013) report that when carbon is taxed at social cost, the price of electricity increases by about 33% relative to its baseline value.[21] These two stylized facts allow me to pin down values for the pollution disutility parameter β simultaneously with choosing a baseline, business-as-usual value for the emissions tax τ. The reason that the emissions tax requires a business-as-usual value, not simply zero, is that the model will not solve with a zero emissions tax, since emissions, modeled as an input, when free will be used in infinite quantity. Thus, I impose a business-as-usual tax rate τ_{BAU} representing the market costs of the energy inputs that lead to carbon emissions

[21] See their Figure 3, comparing the "High" carbon tax simulation (which most closely corresponds to the optimal carbon tax reported for 2050 in Barrage and Nordhaus (2024)) with the baseline simulation. Though their simulation extends only to 2035, not to 2050, this price increase is steady across all years of their simulation.

(i.e. fossil fuels). I calibrate this tax rate and β simultaneously so that the optimal emissions tax rate is 33% higher than the business-as-usual tax (energy price), and so that the optimal emissions level is 30% lower than business-as-usual. This yields $\beta = 0.3$ and $\tau_{BAU} = 0.5135$ (both in arbitrary units).

When the number of consumers $N = 5$, under these parameter values, the optimal level of emissions solving the planner's problem is $E = 0.1278$ and the optimal Pigouvian pollution tax that implements this level of pollution is $\tau = 0.6830$ (when the dirty good is the numeraire so that $p_x = 1$). These values are in arbitrary model units, so I convert them to real-world units based on the calibration. Since Barrage and Nordhaus (2024) report that the optimal carbon tax in 2050 is $125 per ton of CO_2, in 2019 dollars, I convert the pollution tax value τ so that the difference between the optimal pollution tax and the business-as-usual tax is equivalent to $125 per ton. I present results on emissions quantities in terms of how large they are relative to the business-as-usual emissions quantity. At the calibrated values, the optimal emissions quantity is about 67% of the business-as-usual quantity (slightly below the calibration target of 70% from Paul et al. (2013)).

A.III. Additional Equity Simulations

All the plots in the main text of the paper keep the exogenous income distribution (described by the θ parameters) fixed, calibrated based on U.S. data as described in Appendix section A.II. Here, I present some additional simulations that vary these parameters to explore how the exogenous income distribution affects optimal policy. In

Appendix Figure A1, I vary them and simulate how emissions in a decentralized economy respond to the Pigouvian emissions tax, which does not consider income inequality. For each of the vectors of parameters describing the income distribution, $\theta_x, \theta_z, \theta_l$, and θ_{tau}, I vary them in the following way. I begin with setting them all equal to constant values (each element equaling $1/N$), representing no income inequality (the left-most point of Appendix Figure A1). I then linearly change each parameter value until it reaches the calibrated value used in the main text; this corresponds to a variance in the θ_x vector of about 0.14. Although I linearly vary the parameters in all the income distribution vectors, the x-axis scale is represented by just the variance in θ_x. Under each set of parameter values, I solve the decentralized model responding to the Pigouvian emissions tax. According to the theory, the Pigouvian emissions tax τ is independent of these ownership shares, since this tax is based on just the planner's problem in which these shares do not appear. Appendix Figure A1 presents the resulting emissions level for each value of the ownership share parameters θ, when the emissions tax is kept constant at the Pigouvian level. The emissions level is not constant with respect to the ownership shares, given this constant Pigouvian tax τ. When the income is split exactly evenly (variance of zero), then the resulting emissions level is identical to that under the planner's problem (about a 66.2% reduction relative to business-as-usual). But when there is higher income inequality (larger variance) then the resulting emissions level is lower. This result arises from the fact that the Pigouvian emissions tax, solved via the planner's problem and thus independent of the income distribution, will only yield the first-best solution when a regulator also has access to lump-sum transfers. In this simulation, the emissions tax is the only instrument, and so

the variation in ownership shares influences the total level of emissions in response to the Pigouvian tax; this variation would be absent in the presence of lump-sum transfers.

Appendix Figures A2 and A3 present solutions for the regulator's problem as the variance of the income distribution is changed in the same way as it was for Appendix Figure A1. That is, in these simulations the regulator is choosing an emissions tax to maximize social welfare given the constraints of the decentralized problem including the exogenous income shares. The emissions tax that solves the regulator's problem, as shown in Appendix Figure A2, varies with the variance in income, decreasing from exactly $125 per ton (the Pigouvian tax rate) when there is no income inequality to about $121 per ton at the highest level of income inequality (which corresponds to the tax rate that solves the regulator's problem under uniform social welfare weighting in Figure 1). Appendix Figure A2 demonstrates that in fact the emissions tax that maximizes social welfare is not independent of the ownership shares. A more unequal distribution of income yields a higher tax rate. But this higher tax rate under the more unequal distribution yields a lower level of emissions, as shown in Appendix Figure A3. Though the tax rate is lower with a higher variance in θ_x, the higher variance in θ_x itself means that this lower tax rate results in slightly lower emissions. In the regulator's problem as I have framed it, the regulator can only choose the emissions tax and has no access to any income transfer, lump-sum or otherwise. Because of this constraint, the emissions tax must be modified from the Pigouvian rate in the presence of inequalities that erode social welfare, all else equal, and the emissions level will not match the first-best emissions level. In unreported simulations, I confirm that when the regulator also has access to a lump-sum transfer, the chosen

pollution tax is always the Pigouvian tax, while the magnitude of the transfer depends on the variance in the income distribution.

A.IV. Alternative Incommensurability Specification

The social welfare function incorporates how consumers value the environment by including the utility function $U_i(x_i, z_i; E)$. As described earlier, the planner can respect that consumers value the environment though not directly include the environment in their social welfare function, according to this approach. But some may argue that this violates the principle of the incommensurability of values. To explore this further, I consider an approach where the social planner only considers consumer utility over the market goods and not over the environment; i.e. the planner uses the following utility function in the social welfare function: $U_i^0(x_i, z_i)$. In principle, this utility function should be identical to the utility function that the consumers employ except for omitting the environmental quality E. In practice, this may be infeasible to implement, for instance if environmental quality interacts with the consumption of one or both of the market goods as a complement or substitute. In my calibration, though, environmental quality enters the utility function $U_i(x_i, z_i; E)$ additively separably from the market goods, so it is straightforward to remove E to create the alternative utility function $U_i^0(x_i, z_i)$. Under this specification, the optimal emissions tax becomes $\tau = -p_x \left[\frac{\lambda_V}{\lambda_x} V'(E)\right]$. The environmental quality constraint always binds, so that $V(E) = \bar{V}$. In these simulations, the level of the emissions tax is increasing in $\bar{V}$ over all values; there is no kink between the binding and

non-binding portions. This means that for low values of $\bar{V}$, the emissions tax will be lower than the Pigouvian value, and the emissions level will be higher.

Appendix References

Abrell, Jan, Sebastian Rausch, and Giacomo A. Schwarz. "How robust is the uniform emissions pricing rule to social equity concerns?." *Journal of Environmental Economics and Management* 92 (2018): 783-814.

Anderson, Elizabeth. 1995. *Value in Ethics and Economics*. Harvard University Press.

Anderson, Elizabeth. 2014. “Social Movements, Experiments in Living, and Moral Progress: Case Studies from Britain’s Abolition of Slavery.” *Lindley Lecture*.

d’Autume, Antoine, John M. Hartwick, and Katheline Schubert. "The zero discounting and maximin optimal paths in a simple model of global warming." *Mathematical Social Sciences* 59, no. 2 (2010): 193-207.

Baranzini, Andrea, Jeroen CJM Van den Bergh, Stefano Carattini, Richard B. Howarth, Emilio Padilla, and Jordi Roca. "Carbon pricing in climate policy: seven reasons, complementary instruments, and political economy considerations." *Wiley Interdisciplinary Reviews: Climate Change* 8, no. 4 (2017): e462.

Barrage, Lint, and William Nordhaus. "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model." *Proceedings of the National Academy of Sciences* 121, no. 13 (2024): e2312030121.

Becker, Robert A. "Intergenerational equity: the capital-environment trade-off." *Journal of Environmental Economics and Management* 9, no. 2 (1982): 165-185.

Berta, N., 2020. Efficiency without optimality: environmental policies and pollution pricing in the late 1960s. *Journal of the History of Economic Thought*, *42*(4), pp.539-562.

Berta, Nathalie. "Emissions trading: Commodification of pollution—From resistance to proliferation." In *The Routledge Handbook of Commodification*, pp. 351-362. Routledge, 2023.

Bruegge, Chris, Tatyana Deryugina, and Erica Myers. "The distributional effects of building energy codes." *Journal of the Association of Environmental and Resource Economists* 6, no. S1 (2019): S95-S127.

Budolfson, Mark, Francis Dennig, Frank Errickson, Simon Feindt, Maddalena Ferranna, Marc Fleurbaey, David Klenert et al. "Climate action with revenue recycling has benefits for poverty, inequality and well-being." *Nature Climate Change* 11, no. 12 (2021): 1111-1116.

Caney, Simon, and Cameron Hepburn. 2011. “Carbon Trading: Unethical, Unjust and Ineffective?” *Royal Institute of Philosophy Supplement* 69 (October):201–34.

Cremer, Helmuth, Firouz Gahvari, and Norbert Ladoux. "Externalities and optimal taxation." *Journal of Public Economics* 70, no. 3 (1998): 343-364.

Cremer, Helmuth, Firouz Gahvari, and Norbert Ladoux. "Environmental taxes with heterogeneous consumers: an application to energy consumption in France." *Journal of Public Economics* 87, no. 12 (2003): 2791-2815.

DeCanio, Stephen J., Charles F. Manski, and Alan H. Sanstad. "Minimax-regret climate policy with deep uncertainty in climate modeling and intergenerational discounting." *Ecological Economics* 201 (2022): 107552.

Drupp, Moritz A., Ulrike Kornek, Jasper N. Meya, and Lutz Sager. "The economics of inequality and the environment." *Journal of Economic Literature* 63, no. 3 (2025): 840-874.

Fleurbaey, Marc, and François Maniquet. "Optimal income taxation theory and principles of fairness." *Journal of Economic Literature* 56, no. 3 (2018): 1029-1079.

Fried, Stephie, Kevin Novan, and William B. Peterman. "Understanding the inequality and welfare impacts of carbon tax policies." *Journal of the Association of Environmental and Resource Economists* 11, no. S1 (2024): S231-S260.

Glaeser, Edward L., Caitlin S. Gorback, and James M. Poterba. "How Regressive are Mobility-Related User Fees and Gasoline Taxes?." *Tax Policy and the Economy* 37, no. 1 (2023): 1-56.

Goodin, Robert E. 1994. “Selling Environmental Indulgences.” *Kyklos* 47 (4): 573–96.

Gowdy, John, and Jon D. Erickson. 2005. “The Approach of Ecological Economics.” *Cambridge Journal of Economics* 29 (2): 207–22.

Hassler, John, Per Krusell, and Conny Olovsson. Energy-saving technical change. No. w18456. National Bureau of Economic Research, 2012.

Hormio, Säde. "Collective responsibility for climate change." *Wiley Interdisciplinary Reviews: Climate Change* 14, no. 4 (2023): e830.

Howe, Roger E., and John E. Roemer. "Rawlsian Justice as the Core of a Game." *The American Economic Review* 71, no. 5 (1981): 880-895.

Jacobs, Bas, and Frederick van der Ploeg. "Redistribution and pollution taxes with non-linear Engel curves." *Journal of Environmental Economics and Management* 95 (2019): 198-226.

Kelleher, J. Paul. *The Social Cost of Carbon: Ethics and the Limits of Climate Change Economics*. Oxford University Press, 2025.

Kim, Hajin. "Does Paying to Pollute Make Pollution Seem Less Bad?." *Available at SSRN* (2024).

Klenert, David, Linus Mattauch, Emmanuel Combet, Ottmar Edenhofer, Cameron Hepburn, Ryan Rafaty, and Nicholas Stern. "Making carbon pricing work for citizens." *Nature Climate Change* 8, no. 8 (2018): 669-677.

Krutilla, John V. 1967. “Conservation Reconsidered.” *The American Economic Review* 57 (4): 777–86.

Levi, Sebastian. "Why hate carbon taxes? Machine learning evidence on the roles of personal responsibility, trust, revenue recycling, and other factors across 23 European countries." *Energy Research & Social Science* 73 (2021): 101883.

Martinez-Alier, Joan, Giuseppe Munda, and John O'Neill. "Weak comparability of values as a foundation for ecological economics." *Ecological economics* 26, no. 3 (1998): 277-286.

Mintz-Woo, Kian. 2022. “Carbon Pricing Ethics.” *Philosophy Compass* 17 (1): e12803.

Obradovich, Nick, and Scott M. Guenther. "Collective responsibility amplifies mitigation behaviors." *Climatic change* 137 (2016): 307-319.

O'Neill, John. 2017. "Pluralism and Incommensurability." In *Routledge Handbook of Ecological Economics*, 227–36. Routledge.

Page, Edward A. 2011. "Cosmopolitanism, Climate Change, and Greenhouse Emissions Trading." *International Theory* 3 (1): 37–69.

Paton, Joy, and Gareth Bryant. "Valuing pollution: problems of price in the commodification of nature." *The Economic and Labour Relations Review* 23, no. 1 (2012): 87-106.

Rawls, John. *A Theory of Justice*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1971

Raz, Joseph. 1985. "Value Incommensurability: Some Preliminaries." *Proceedings of the Aristotelian Society* 86:117–34.

Sandmo, Agnar. "Optimal taxation in the presence of externalities." *The Swedish Journal of Economics* (1975): 86-98.

Smessaert, Jacob, Antoine Missemer, and Harold Levrel. "The commodification of nature, a review in social sciences." *Ecological Economics* 172 (2020): 106624.

Thoma, Johanna, 2024. Social Science, Policy and Democracy. *Philosophy & Public Affairs*, *52*(1), pp.5-41.

van der Ploeg, Frederick, Armon Rezai, and Miguel Tovar Reaños. "Third-Best Carbon Taxation: Trading off emission cuts, equity, and efficiency." *Journal of the Association of Environmental and Resource Economists* 12, no. 4 (2025): 1023-1058.

Vanderheiden, Steve. "Climate change and collective responsibility." In *Moral responsibility: Beyond free will and determinism*, pp. 201-218. Dordrecht: Springer Netherlands, 2011.

Walraevens, Benoît. "Rawls's maximin and optimal taxation theory." *The European Journal of the History of Economic Thought* 30, no. 5 (2023): 860-882.

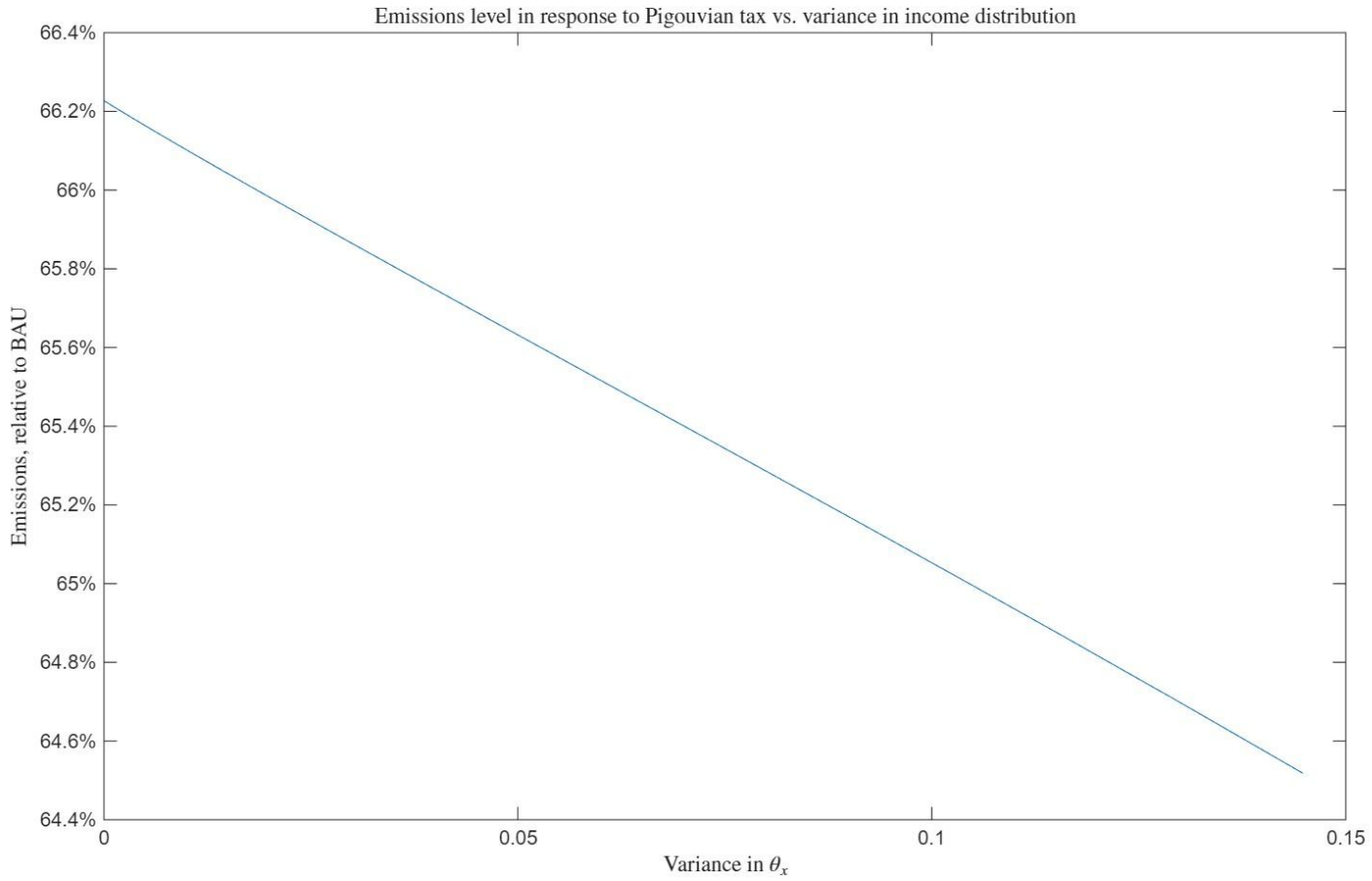
Wiese, Catharina, Richard Cowart, and Jan Rosenow. "The strategic use of auctioning revenues to foster energy efficiency: status quo and potential within the European Union Emissions Trading System." *Energy efficiency* 13, no. 8 (2020): 1677-1688.

Appendix Table A1

Para-meter	Description	Calibration/Source	Value
N	Number of consumers		5
θ_{ix}	Income share of dirty firm (x) revenue across consumers	CBO (2008), Table 5	$[.01, .02, .03, .06, .87]$
θ_{iz}	Income share of clean firm (z) revenue across consumers	CBO (2008), Table 5	$[.01, .02, .03, .06, .87]$
θ_{il}	Income share of wage income across consumers	CBO (2008), Table 5	$[.03, .09, .15, .23, .50]$
$\theta_{i\tau}$	Income share of pollution tax revenue across consumers	Impose uniform return ($\theta_{i\tau} = 1/N$)	$[.2, .2, .2, .2, .2]$
κ	Labor share in dirty good production (Cobb-Douglas)	Hassler et al. (2012)	0.7
μ	Exponent in clean good production	Hassler et al. (2012)	0.7
α	Utility power for dirty good (x)	Ensure interior solution	0.3
ω	Utility power for clean good (z)	Ensure interior solution	0.4
β	Disutility of pollution	Target so that first-best emissions is 30% lower than BAU emissions	0.3

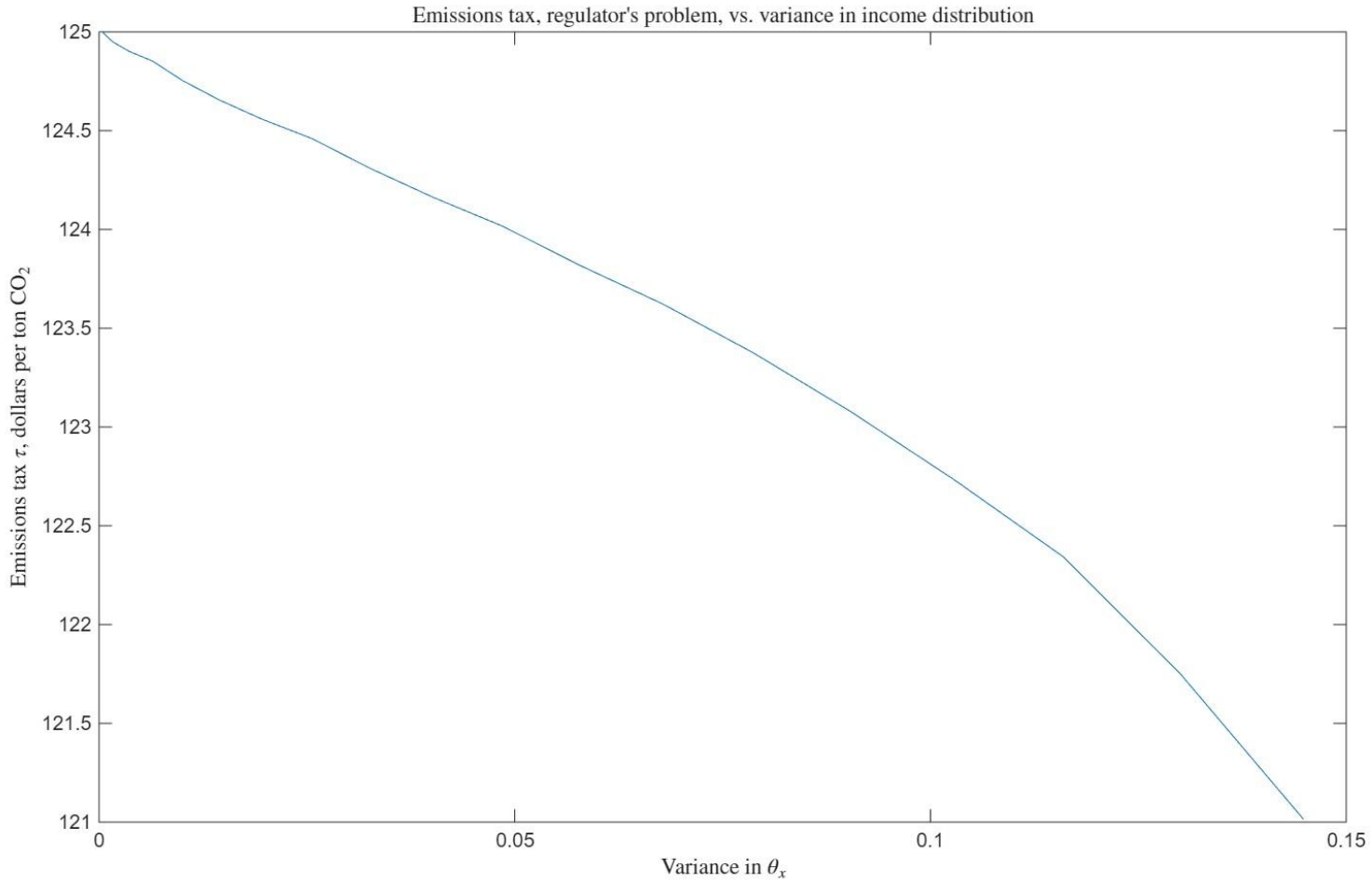
Notes: This table lists the parameter values used in the base case of the numerical simulations along with their sources or strategies for calibration.

Appendix Figure A1



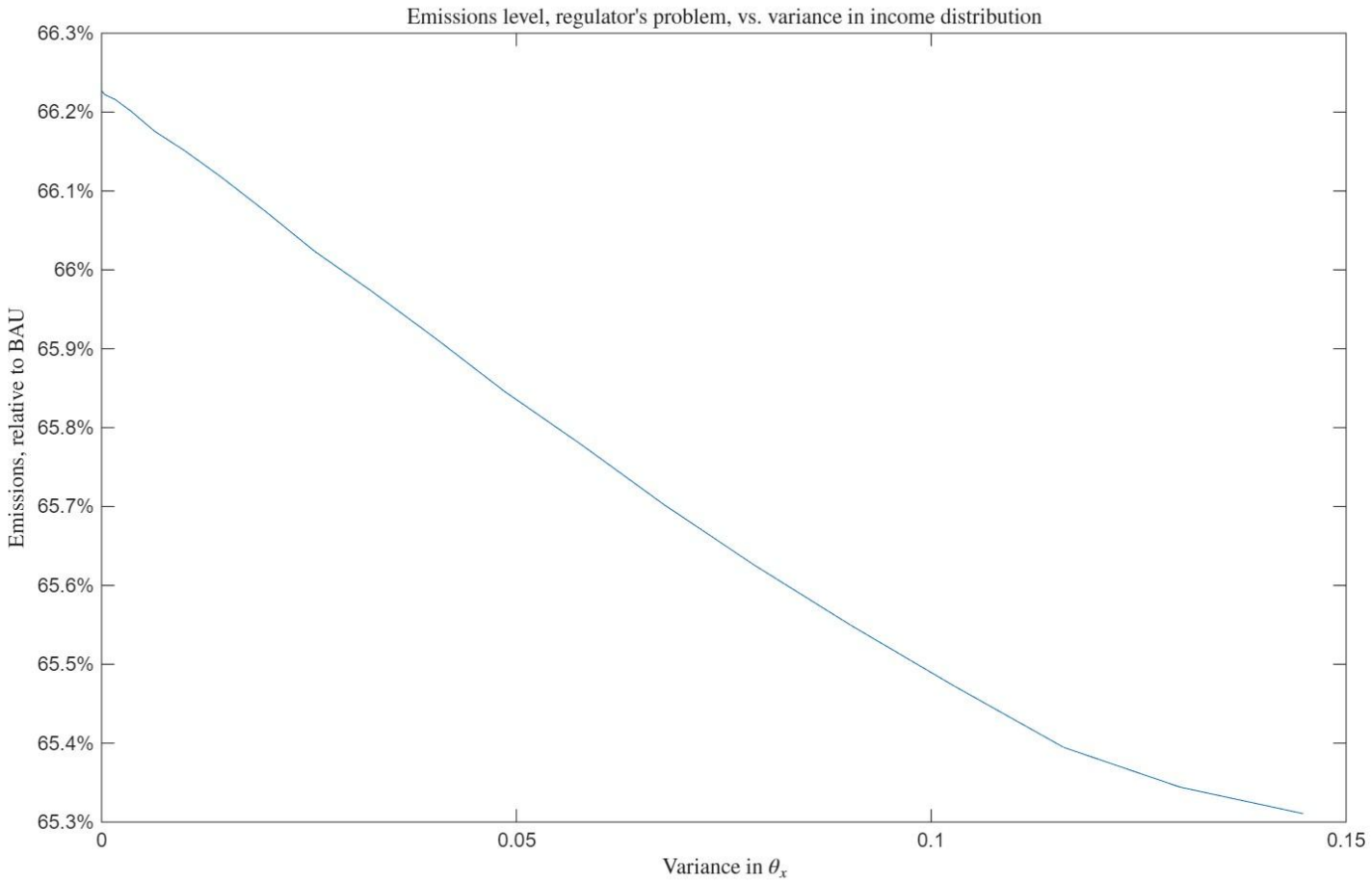
Notes: The x-axis is the variance in θ_x, the vector of the shares of the polluting firm's profits earned by each representative consumer. It ranges from 0, with no income inequality, to about 0.14, which corresponds to the baseline calibrated values. The variance in the shares of the other income sources ($\theta_z, \theta_l, \theta_{tau}$) are varied accordingly. The y-axis is the quantity of emissions that results from the decentralized economy in response to a pollution tax kept at the fixed Pigouvian rate, expressed as a percentage of the business-as-usual emissions level.

Appendix Figure A2



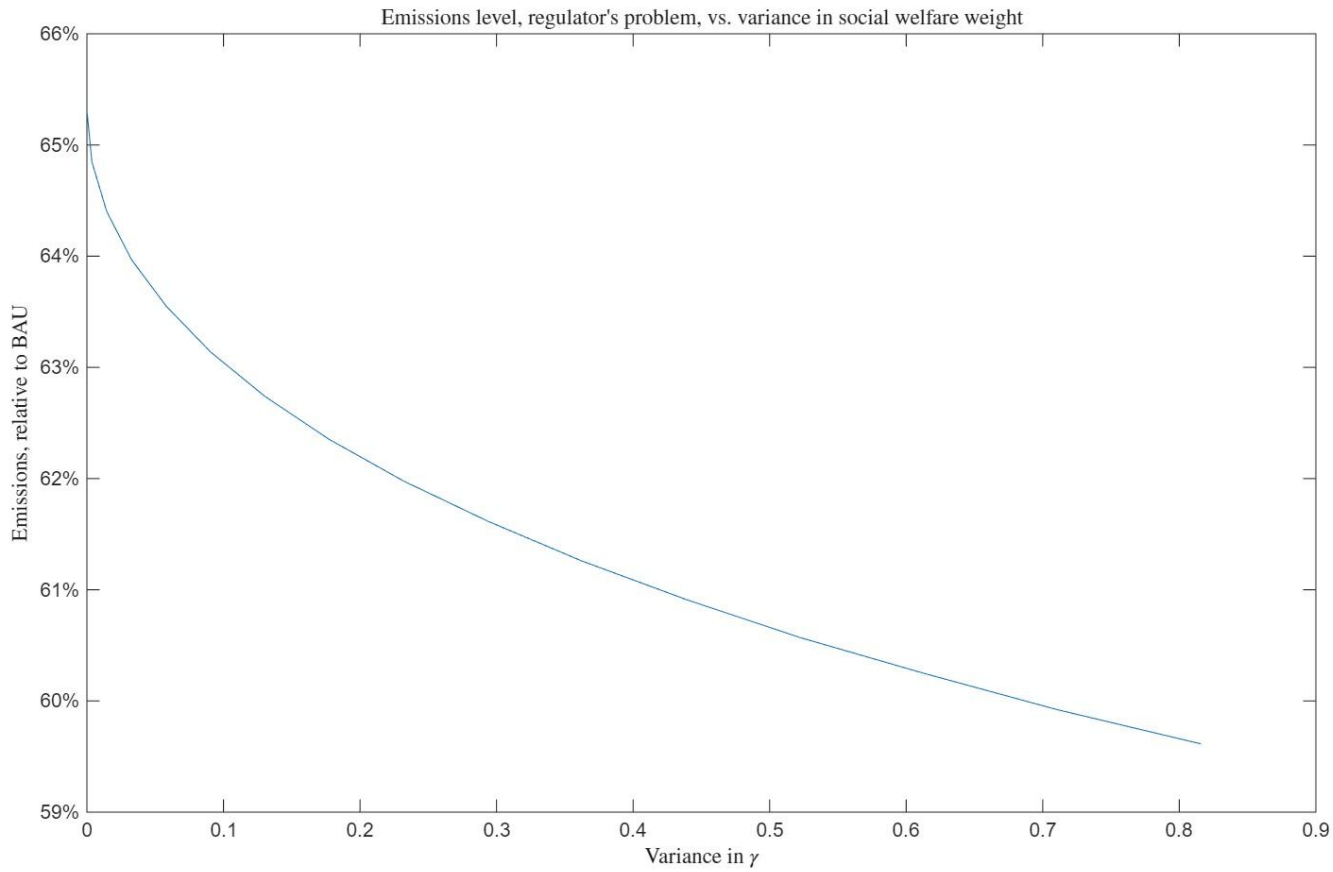
Notes: The x-axis is the variance in θ_x, the vector of the shares of the polluting firm's profits earned by each representative consumer. It ranges from 0, with no income inequality, to about 0.14, which corresponds to the baseline calibrated values. The variance in the shares of the other income sources ($\theta_z, \theta_l, \theta_{tau}$) are varied accordingly. The y-axis is the emissions tax rate that solves the regulator's problem, in 2019 dollars per ton of carbon dioxide.

Appendix Figure A3



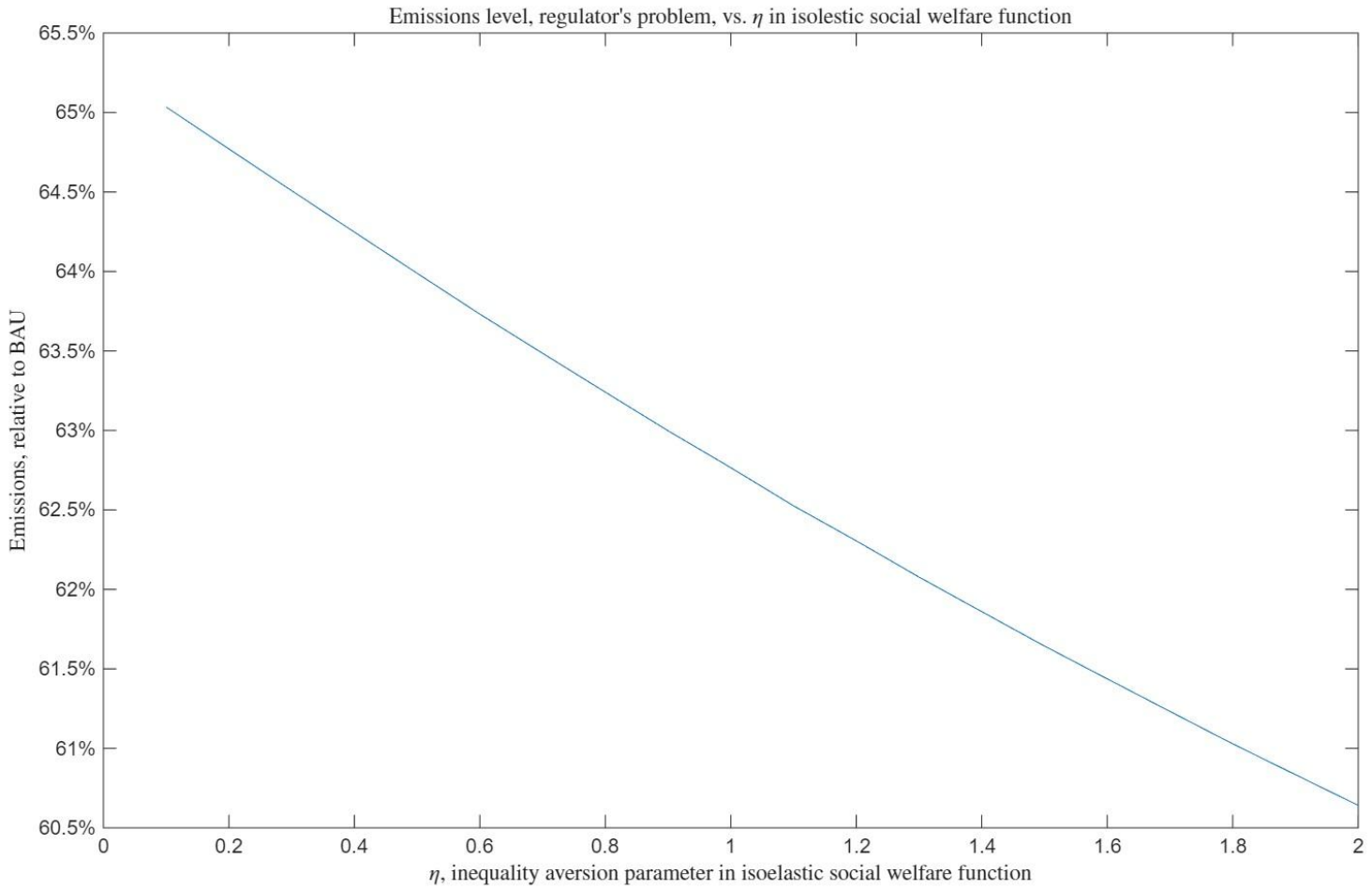
Notes: The x-axis is the variance in θ_x, the vector of the shares of the polluting firm's profits earned by each representative consumer. It ranges from 0, with no income inequality, to about 0.14, which corresponds to the baseline calibrated values. The variance in the shares of the other income sources ($\theta_z, \theta_l, \theta_{tau}$) are varied accordingly. The y-axis is the quantity of emissions that results from the solution to the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Appendix Figure A4



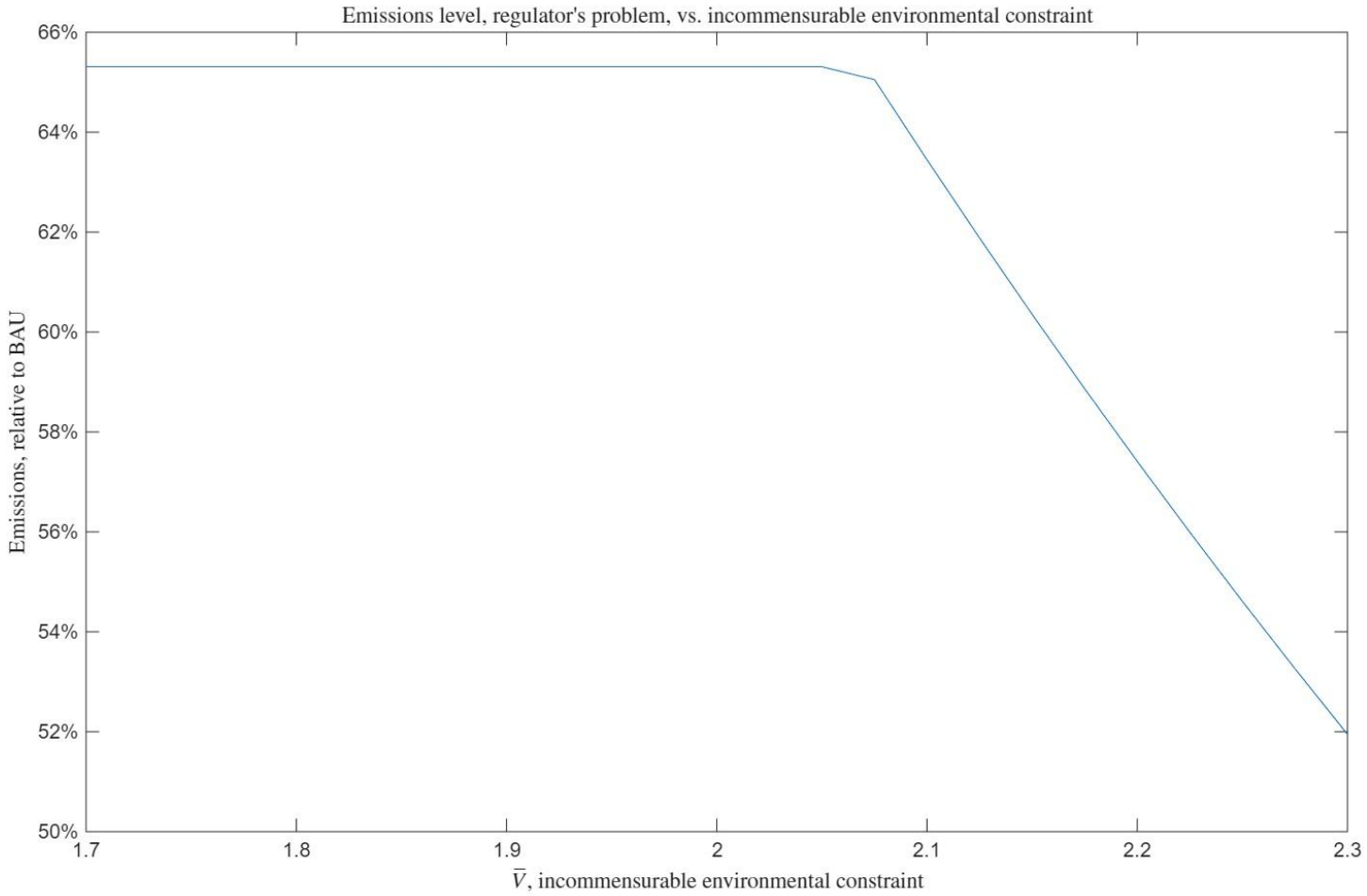
Notes: The x-axis is the variance in the vector of social welfare weights γ, ranging from zero (where all the γ_i values equal one) to 0.82 (where the γ_i for the richest quintile is 0.25 and the γ_i for the poorest quintile is 2.5). The y-axis is the quantity of emissions that results from the solution to the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Appendix Figure A5



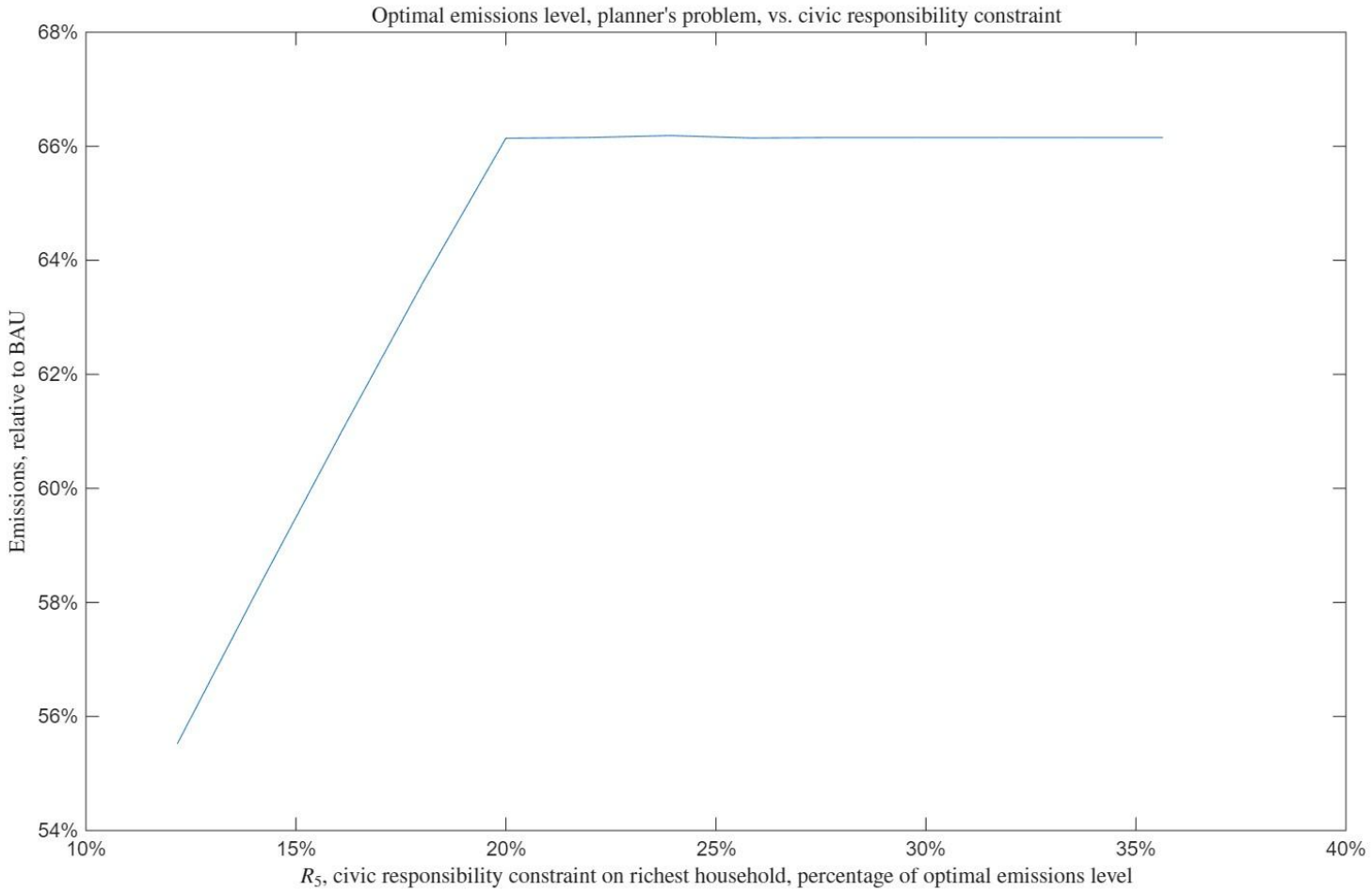
Note: The x-axis is η, the inequality aversion parameter in the isoelastic social welfare function. The y-axis is the quantity of emissions that results from the solution to the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Appendix Figure A6



Notes: The x-axis is $\bar{V}$, the value of the minimum environmental quality constraint in the incommensurability specification. The y-axis is the optimal quantity of emissions from the regulator's problem, expressed as a percentage of the business-as-usual emissions level.

Appendix Figure A7



Notes: The x-axis is R_5, the value of the civic responsibility on the richest household, expressed as a percentage of the overall optimal emissions level (solving the planner's problem). The constraints on the other four households are kept constant at 20% of the overall optimal emissions level. The y-axis is the optimal quantity of emissions from the planner's problem, expressed as a percentage of the business-as-usual emissions level.