DUKE ENVIRONMENTAL AND ENERGY ECONOMICS WORKING PAPER SERIES
organized by the
NICHOLAS INSTITUTE FOR ENVIRONMENTAL POLICY SOLUTIONS
and the
DUKE UNIVERSITY ENERGY INITIATIVE

Horizontal Equity Effects in Energy Regulation

Carolyn Fischer[*]
William A. Pizer[§]

Working Paper EE 18-01
July 2018

[*] Resources for the Future, Vrije Universiteit Amsterdam, Gothenburg University, Fondazione Eni Enrico Mattei

[§] Sanford School and Nicholas Institute for Environmental Policy Solutions, Duke University; Resources for the Future; National Bureau of Economic Research

Acknowledgments

Carolyn Fischer is grateful for the support of the European Community's Marie Skłodowska–Curie International Incoming Fellowship, "STRATECHPOL – Strategic Clean Technology Policies for Climate Change," financed under the EC Grant Agreement PIIF-GA-2013-623783. This work was supported by the Alfred P. Sloan Foundation, Grant G-2016-20166028.

Horizontal Equity Effects in Energy Regulation

Carolyn Fischer and William A. Pizer*

Abstract

Choices in energy regulation, particularly whether and how to price externalities, can have widely different distributional consequences both across and within income groups. Traditional welfare theory focuses largely on effects across income groups; such "vertical equity" concerns can typically be addressed by a progressive redistribution of emissions revenues. In this paper, we review alternative economic perspectives that give rise to equity concerns within income groups, or "horizontal equity," and suggest operational measures. We then apply those measures to a stylized model of pollution regulation in the electricity sector. In addition, we look for ways to present the information behind those measures directly to stakeholders. We show how horizontal equity concerns might overshadow efficiency concerns in this context.

Key Words: equity, inequality, cap-and-trade, carbon price, performance standards

JEL Classification Numbers: D61, D63, Q48, Q52, Q58

* Carolyn Fischer is Senior Fellow at Resources for the Future (fischer@rff.org); Professor at Vrije Universiteit Amsterdam; Marks Visiting Professor at Gothenburg University; Visiting Senior Researcher at Fondazione Eni Enrico Mattei; and CESifo Research Network Fellow. William A. Pizer is Susan B. King Professor of Public Policy, Sanford School, and Faculty Fellow, Nicholas Institute for Environmental Policy Solutions, Duke University (william.pizer@duke.edu); University Fellow, Resources for the Future; Research Associate, National Bureau of Economic Research. Fischer is grateful for the support of the European Community's Marie Skłodowska–Curie International Incoming Fellowship, "STRATECHPOL – Strategic Clean Technology Policies for Climate Change," financed under the EC Grant Agreement PIIF-GA-2013-623783. This work was supported by the Alfred P. Sloan Foundation, Grant G-2016-20166028.

Introduction

Economists often give primacy to the efficiency or cost-effectiveness of regulatory design, favoring Pigouvian pricing mechanisms for addressing environmental externalities. Implicitly or explicitly, economists' favoritism assumes equity concerns can be dealt with by allocating the rents created by emissions pricing. For example, tax rate changes can redistribute rents to achieve a desired level of progressivity with respect to impacts across income group, often with particular attention to about outcomes for poor households.

In this paper, we make three observations that together suggest favoritism for Pigouvian policies may be misplaced because equity concerns are not so easily vanquished. First, the focus on equity as a question of impacts across income group, and the poor as a collective group, derives from traditional welfare theory. Traditional welfare theory places value on equalizing household utility and bringing the poor and rich closer together. However, an alternative line of thought, typically referred to as fair burden or horizontal equity (HE), places value on similar households facing similar changes. The distribution of impacts across income groups still matters, but so do impacts within income groups.

Second, Pigouvian pricing policies can involve household-level costs and benefits that are orders of magnitude larger than those arising under other, non-Pigouvian policies. This is true even as positive and negative household-level impacts cancel out in the aggregate and Pigouvian policies are less expensive for society. Finally, the redistribution created by Pigouvian pricing of energy externalities can be substantially unrelated to income and other easily observable variables. This makes it difficult if not impossible to neutralize large, unequal effects within income groups.

Taken together, these observations suggest that Pigouvian energy regulation may have relatively large, unavoidable, horizontal equity impacts and that the economists' favoritism may be misplaced. It also raises a question: How can policy analysts present information about horizontal equity in ways that facilitate stakeholder discussion and policymaker decisions?

We note this is not always the case. Some Pigouvian energy regulation does not directly affect households and equity impacts are manageable. Under the acid rain program, for example, an agreement was made to cap sulfur dioxide (SO_2) emissions from coal-fired power plants at a specific level. With a general notion of how to allocate emissions rights, attention then shifted to horse-trading among the companies to address the exact distribution of burden (Cohen 1995). But, importantly, the price of electricity was largely unaffected (Burtraw et al. 2005). Natural gas generators are often the marginal producers and do not emit SO_2. Hence, power generation companies were the ones who felt the effects of the regulation and allocation choices—and were directly involved in those choices. Coasean bargaining at its best.

Market-based CO_2 programs, however, have the potential to raise electricity and other energy prices significantly. Over a hundred million households, as well as businesses, will feel the direct effect of regulation as well as the choices about allowance or revenue allocations. Direct horse-trading to address equity is difficult if not impossible. Individual bargaining is replaced by generic rules, perhaps based on income or other observable demographics. As we show, it will be difficult to alleviate substantial inequity from energy price impacts based on observable demographics.

While our contribution ties energy regulation to policymaking concerns about horizontal equity, we are not the first paper to remark on these additional distributional impacts of Pigouvian energy pricing. Burtraw and Palmer (2008) find that a carbon pricing policy has net social costs of roughly \$0.5 billion annually, while consumers and producers lose more than \$21 billion in pollution payments. With an eye towards fuel taxes, Poterba (1991) presents gasoline expenditures by income decile but also reports the fraction of each decile both where expenditures shares are above 0.1 and where they are equal to zero. He finds an average share of 0.039 for the lowest decile masks but 36 percent of this decile spend nothing on gasoline while 14 percent have an expenditure share exceeding 0.1. Among practitioners, analysis of tax reform proposals regularly focus on the coefficient of variation of impacts within income groups (Westort and Wagner 2002).

Taking a more expansive approach, Rausch et al. (2011) use graphical figures to present the distributional effects of carbon pricing associated with various rebate approaches. They present box-and-whisker plots, similar to our preferred graphical figures, showing outcomes across and within income deciles. They highlight that some amount of progressivity and regressivity is certainly present, with the mean cost by decile ranging from 0 to 0.5 percent of income. At the same time, a large number of households experience gains and losses of more than 1 percent.

None of these papers, however, suggests that there might be a welfare cost to substantial variation in household effects within income groups. Only Burtraw and Palmer (2008) point out that non-Pigouvian policies can lead to much smaller distributional effects. In this way, we believe our paper offers a new perspective on efficiency-equity concerns.

In order to make our points about horizontal equity and Pigouvian pricing, we first review the various rationales for valuing horizontal equity as well as the controversies. We then present two welfare measures to operationalize these ideas and explain how they relate to models in the literature. In order to relate this to energy policy, we then turn to how energy regulation affects energy prices and ultimately household welfare. We consider a stylized model of two climate change policies. One policy is Pigouvian pricing, a mass-based cap and trade (CAT) policy applied to the electric power sector with auction revenue used to provide an equal rebate per household. The other is non-Pigouvian, a rate-based tradeable performance standard (TPS) that is effectively a revenue-neutral combination of a tax on emissions and subsidy to output of electricity. Thanks to the subsidy, the TPS does not raise electricity prices as much as does the CAT.

To complete the task, we simulate the distributional consequences of these policies using a sample of observed consumer expenditures. Within this sample we see a large heterogeneity of household electricity expenditures even within a single income group. The CAT therefore results in much more horizontal, within-income-group redistribution than the TPS. We put these outcomes into our welfare measure and show that this can translate into lower welfare under CAT versus TPS. Finally, we consider ways that one might present this information to stakeholders and policymakers without appealing to elaborate welfare theories, but while still remaining consistent with those theories. And we revisit the degree to which horizontal equity might be attenuated using other observable data, arguing that such efforts are unlikely to help.

Foundations of horizontal equity in economic thought

Equity and justice have long been principles in public economics (Elkins 2006). Within this rich intellectual history, we identify two threads that speak to the idea of treating similar households similarly in public policy. There is an older literature that frames the discussion in terms of *equal sacrifice* regarding the provision of public goods and a more recent, welfarist approach that builds on the axiomatic treatment of welfare measures. The latter encompasses both provision of public goods and redistribution from rich to poor. Beyond public economics, one can interpret the behavioral work by Tversky, Kahneman, and others, as supporting horizontal equity. Here, we review these ideas before turning to operationalizing our approach.

Equal Sacrifice

The principle of equal sacrifice dates at least to the 19th century. For John Stuart Mill (1871), "Equality of taxation ... means equality of sacrifice. It means apportioning the contribution of each person toward the expenses of government so that he shall feel neither more nor less inconvenience from his share of the payment than every other person experiences from his." This principle of equal sacrifice in paying for public goods could be interpreted as supporting progressive taxation, to ensure equal consequences in terms of utility and to ensure that equally situated persons are treated equally. The 19th-century utilitarian philosopher and economist Henry Sidgwick considered equal sacrifice the "obviously equitable principle—assuming that the existing distribution of wealth is accepted as just or not unjust" (Weinzierl 2012). In other words, assuming society does not want to engage in additional income redistribution, the burdens of financing government should be shared equally.

The question of whether society does or does not want to engage in income redistribution is an important distinction. This older literature tended to separate this question from the question of how to fund public goods. Tracing back to the Greeks, Elkins (2006) argues that the principal of equal treatment can be seen as an application of Aristotelian philosophy, in which a just distribution is based on merit. If individuals "merit" their status quo ante distribution, then they should merit equal shares in the post-intervention distribution. Of course, as the definition of merit matters for evaluating the fairness of Aristotelian justice, "the moral basis of horizontal equity depends upon the moral standing of the market distribution" (p. 73).

A second distinction framed in the early literature is the fair treatment of similar individuals versus the fair treatment of very different individuals. In a treatise on tax policy, Simons (1938) states that "taxes should bear similarly upon persons similarly situated" (p. 106). Pigou himself noted that "equal sacrifice among similar and similarly situated persons is an entirely different thing from equal sacrifice among all persons" (Pigou 1928).

Welfarism

While the distinction between vertical and horizontal equity came to prominence in the early and mid-20th century, it was the welfarists in the latter part of the 20th century that introduced these terms. In that context, mitigating social inequality became referred to as "vertical equity," while treating people in similar circumstances similarly became recognized as "horizontal equity" (Elkins 2006, 43). Now, in addition to questions about taxes and the provision of public good, social policy explicitly considered redistribution.

Welfarism looks at the desirability of public policy in terms of whether the state of affairs with the policy has a higher welfare measure than the state without (Sen 1979). A distinctive feature of horizontal equity is the comparison to a reference point, rooted in the *status quo ante*. Whereas vertical equity can be measured for any distribution of income or utility (such as with a Gini coefficient, before or after a policy intervention), assessing horizontal equity requires a change to evaluate. Similarly situated persons are so situated *ex ante*, and that reference point sits within a pre-existing vertical distribution. Most traditional, axiomatic welfare measures (see

Chapter 23 of Mueller 2003) avoid reference points and fail to capture horizontal equity. Moreover, it is inclusion of a reference point that has regularly led to controversy.

Different approaches have been taken with respect to reference points. Early economic applications of horizontal equity in public finance focused on rank as a reference point. The *Pigou-Dalton axiom* holds that a social welfare function should prefer allocations that are more (vertically) equitable, as long as redistribution does not change the ranking of individuals. Adler (2013, p. 1) defends this "prioritarian" view, adjusting the measure of well-being according to responsibility: "if one person is at a higher level of well-being than a second, and the worse-off one is not responsible for being worse off, then distributive justice recommends a non-leaky, non-rank-switching transfer of well-being from the first to the second, if no one else's well-being changes."

A later application by Auerbach and Hassett (2002) introduces reference points by nesting groups with similar, pre-policy incomes into an aggregate welfare function. Separate elasticity parameters penalize income inequality within the nested groups (horizontal equity) and across nested groups (vertical equity). Otherwise, the aggregate function looks like a more traditional welfare function of post-policy outcomes, not changes. In practice, this makes it difficult to measure horizontal equity effects arising from a new policy when changes in income are small relative to existing differences between individuals within each nested group.[1] This is often the case with energy regulation.

In a series of articles, Louis Kaplow (1989, 1992, 2000) critiques both applications and underlying principles of horizontal equity. He is particularly critical of operationalizing the early

[1] Auerbach and Hassett remark that the horizontal equity effects of income taxes are equivalent to an across the board 0.2-0.4 percent tax increase—roughly 0.01 of total tax costs.

focus on rank, where large rank-preserving redistributions would have to be compared to infinitesimal rank-inverting redistributions. The implied idea of large discontinuities in a welfare measure is unappealing. But he is more generally critical of the notion of a valid refence point. He argues it is counter to the idea of economic mobility. He suggests that the status quo as a reference point arbitrarily treats policy outcomes as more significant than the "luck" leading to status quo differences. Moreover, once a policy is implemented it becomes the status quo. If a policy has negative HE consequences, so does reversing the policy. He points out that HE implies a trade-off with the Pareto principal. Even if no one is worse off, there may be unequal treatment of similar households (though we show this need not alter Pareto welfare rankings). While it is beyond the scope of this paper to respond to all Kaplow's criticisms—criticisms that we view as pointing out logistical consequences but not being fatal to the idea of HE in any case—there is certainly evidence that people think in terms of reference points.

Behavioral Economics

Distinct from the philosophical origins of horizontal equity, behavioral economics provides another motivation for believing reference points are important. In particular, theoretical foundations for reference-based utility were offered by psychologists Kahneman and Tversky (1979), who propose *prospect theory* as a way to incorporate observed behavioral biases in decisionmaking. Central concepts are that people evaluate outcomes relative to a reference point, and gains are evaluated differently from losses, expressed by "loss aversion." Kahneman and Tversky were not explicit about the origin of the reference point, but proposed candidates have been the expected outcome (Kőszegi and Rabin 2006, 2007, 2009), the status quo (the "endowment effect" in Thaler 1980), or the average outcome of others. Although prospect theory was postulated for decisionmaking under uncertainty (and also includes concepts related to biases in evaluating high-risk, low-probability events), Michaelson (2015, p. 202) argues that the same biases also hold for resource distribution problems in the aggregate. His findings "suggest that neither utilitarian nor Rawlsian objectives will properly describe what most people believe is fair." Thus, reference-point biases offer additional support for considering aspects of horizontal equity in policymaking.

Applications to Energy and Environmental Policy

There is reason to believe that horizontal equity issues can loom much larger than vertical equity ones for environmental policy. First, broad-based tax policy is the government's primary tool for addressing vertical inequality; environmental policy is an indirect one at best. If one believes that the overall tax system has evolved to address social inequality to the extent that the existing distribution is "just," then a reasonable equity principle for allocating the burden of environmental policy is to avoid distortions to that distribution. That is, equal sacrifice relative to the status quo. Second, environmental policy costs tend to be small compared to income and other taxes but highly heterogeneous. In the application we consider, impacts are on the order of tens or hundreds of dollars per household.[2] Such changes are unlikely to impact vertical equity in a meaningful way. Nonetheless, equity and fairness concerns remain in the same way that overall cost-benefit concerns remain. And, while most households have options to change behavior and reduce their energy consumption, some margins may be constrained by housing and climate conditions, family size, landlord-tenant relationships, etc. These constraints may

[2] The threshold for "significant regulatory action" requiring cost-benefit anlaysis is $100 million, or just $1 per household (US Government 1993).

need to be taken into account in assessing equity concerns (e.g., "responsibility" and "merit" in the prioritarian and Aristotelian senses).

In the next section, we propose a framework for considering horizontal equity impacts in assessing the costs of a policy and explain its roots. We then show how this relates to the welfare function in Slesnick (1989), which uses changes in utility relative to a reference point, as well as an aggregate welfare function based on the value functions put forward by Tversky and Kahneman (1992).

Operationalizing horizontal equity in welfare theory

As previously discussed, operationalizing horizontal equity into a welfare function faces the challenges of incorporating the referential nature of equal sacrifice while retaining sensible notions regarding redistributions. In this section, we draw on work by Slesnick (1989) and Kahneman and Tversky (1979) to motivate a particular welfare measure that includes horizontal and vertical equity. To make concepts clear, we specify an initial distribution of incomes, $\{y_i^0\}$, for a group of *N* households. These households are affected by a policy that leads to a distribution of changes (consumption variation) given by $\{\Delta y_i\}$. We thus focus on motivating a welfare measure for a specific, policy-induced change in net income:[3]

$$\boldsymbol{W_0 = \overline{\Delta y} - \gamma N^{-1} \textstyle\sum_i |\Delta y_i - r_i|} \tag{1}$$

as well as a slight variant:

[3] The purpose of the measures is to facilitate an evaluation of net benefits from a single policy or a choice among policies. We intentionally refer to W as a "welfare measure for a specific, policy-induced change" without suggesting the measure should be treated as changes in some welfare level that cumulates policy after policy. E.g., we do not refer to our expressions as ΔW.

$$W_1 = \overline{\Delta y} - \gamma \sqrt{N^{-1} \sum_i \frac{\overline{y}}{y_i^0} (\Delta y_i - r_i)^2} \tag{2}$$

In both cases, W is scaled to household, monetary terms in the ballpark of the average household net income change, $\overline{\Delta y} = N^{-1} \sum_i \Delta y_i$. Here, r_i is a reference point for household i, where r_i is constructed such that $N^{-1} \sum_i r_i = \overline{\Delta y}$. The measure W_0 depends on the average absolute deviations from r_i while W_1 depends on the squared (weighted) deviations from r_i. The parameter γ is a weight ($1 \geq \gamma \geq 0$) placed on the second term. We will go through the origins of these welfare function momentarily—importantly, what might generate r_i—but for a moment we highlight a few features.

First, we refer to the second term, after γ, as an "equity penalty" arising from deviations from fair burden (when we want to refer to the term including γ, we will refer to the "weighted equity penalty"). The first term ($\overline{\Delta y}$) measures non-equity costs or benefits, and simply depends on average (or, multiplied by N, total) costs or benefits. This term is unaffected by how those costs or benefits are distributed. The second term measures the effect of deviations from a particular distribution of burden given by the r_i's. These r_i's are able to capture the idea of vertical equity or fairness—the burden that households in different situations ought to bear to achieve the fairest possible outcome. To the extent actual household costs match those defined by the r_i, the penalty is zero. To the extent household costs differ from r, in either vertical (across initially different households) or horizontal (among initially similar households) ways, the penalty is positive and subtracts from welfare.[4]

[4] It is worth noting that this welfare measure is subject to the Kaplow criticism that enacting a policy and then removing it can both involve adverse equity penalties. Imagine a policy that matches fair burden, but then adds some random transfers among similarly situated individuals. The adverse equity penalty that would arise from both implementing and reversing the policy could be viewed as friction.

Second, the penalty is weighted by a scaling factor γ. This is a social choice about the importance of equity concerns and is unavoidable to fully operationalize the welfare metric. As we discuss below, it is natural to constrain γ. In particular, it might not be so large so that making a single person better off, without harming any other, lowers welfare, at least from the status quo, leading to our constraint that $\gamma \leq 1$. That addresses one of Kaplow's criticism, noted above, that HE alone implies a violation of the Pareto principle.

Third, and perhaps most usefully, notions of horizontal and vertical equity can be decomposed. For example, suppose we define the reference point r_i to be the average burden in one's own decile $\overline{\Delta y}_{d(i)}$ where $d(i)$ maps individuals into deciles. (We do this formally in our modeling application section). The penalty now approximates HE only. That is, the deciles as a whole are not penalized for whatever burden they bear, on average. The only penalty is for variation within the decile—whether similar individuals are treated similarly, or not.[5] The additional penalty associated with alternative definitions of r_i comes from vertical inequity.

The fourth and last feature, which we demonstrate in our application, is that our measures allow quick and easy calculation of this HE component of the penalty term based on decile summary statistics. We can construct the HE penalty in W_0 from the average absolute deviation of burden by decile. And the HE penalty in W_1 is approximated based on the standard deviation of burden by decile. That is, if we compute the average absolute deviation of Δy_i and the standard deviation of Δy_i for each decile, the HE penalties in W_0 and W_1 are the simple and (approximately) weighted quadratic average across deciles, respectively, of these two statistics.

[5] We could instead imagine a more complicated scheme that would define r_i in terms of a more localized mean of the Δy_i's, rather than grouping households into deciles. This would be a more precise HE-only measure.

These penalty functions are two cases of a more general penalty function, $\left(N^{-1}\sum_i \left(\frac{u_i'}{\overline{u'}}\right)^{\rho} |\Delta y_i - r_i|^{1+\rho}\right)^{\frac{1}{1+\rho}}$ where u is a utility function and $\rho \geq 0$ is an inequality aversion parameter, which we discuss below.

We now turn to the literature to understand the underlying justification for (1) and (2).

Slesnick

Slesnick (1989) provides the main motivation for our welfare measure. He uses a welfare function based on deviations in household utility u from an initial reference point. Here, we have simplified his model to match our notation, making utility u solely a function of income. Specifically, the change in utility for individual i is given by $\Delta u_i = u(y_i^0 + \Delta y_i) - u(y_i^0)$.

The welfare function begins with a weighted average of utility changes across households, from which is subtracted a measure of deviations from this average. In this way, variation across households in their utility change is costly in terms of welfare, and the welfare-maximizing policy would generally involve an equal utility change across all households. This is the equal sacrifice notion. Slesnick's welfare function can be written as

$$W_s = \overline{\Delta u} - \gamma\left(\sum a_i \left|\Delta u_i - \overline{\Delta u}\right|^{1+\rho}\right)^{\frac{1}{1+\rho}} \tag{3}$$

where $\overline{\Delta u} = \sum_i a_i \Delta u_i$ and $\sum_i a_i = 1$.

We can already see that (3) is somewhat similar to (1) and (2) in functional form with one term capturing the average utility effect and the second a penalty for unequal distribution. That is, the welfare function is increasing in the average utility change but decreasing in a measure of deviations of changes in individual utility from the average. This equity penalty includes horizontal inequity, when individuals with similar incomes face different utility changes. But it also includes vertical inequity, when, collectively, those individuals at a given income level deviate from the income change implied by $\overline{\Delta u}$ at that income level.

Without defining the weights in (3), rearranging costs to minimize deviations in utility changes may affect average utility. However, by weighting the individual deviations by the inverse of marginal utility, we can completely disentangle total costs and burden sharing. Let a_i represent normalized Negishi weights, so $a_i = \frac{u'(y_i^0)^{-1}}{\sum_j u'(y_j^0)^{-1}}$.[6]

When these weights are used—and assuming income changes are small relative to total income, so $\Delta u_i = u'(y_i^0)\Delta y_i$—the average utility change reduces to a rescaled average income change:

$$\overline{\Delta u} = \sum a_i \Delta u_i = \frac{\sum_i u'(y_i^0)^{-1}\Delta u_i}{\sum_i u'(y_i^0)^{-1}} = \frac{\sum_i u'(y_i^0)^{-1} u'(y_i^0)\Delta y_i}{\sum_i u'(y_i^0)^{-1}} = \overline{u'} * \overline{\Delta y}$$

where, as before, $\overline{\Delta y} = N^{-1}\sum \Delta y_i$ is the simple average change in income and $\overline{u'} = (N^{-1}\sum_i u'(y_i^0)^{-1})^{-1}$ is the harmonic average of individual marginal utility. In this way, we see that the first term in (3), $\overline{\Delta u}$, only depends on the average income change, not the how income changes are allocated. That is, we can reallocate dollar costs across households without affecting the first term (or the basis of fair burden in the second term). The penalty is then minimized and

[6] Negishi (1960) formalized an insight for evaluating policies that do not have a primary goal of manipulating the distribution of income. It involved weighting individual utilities by the inverse of the marginal utility of income. With this weighting, the summed social welfare function replicates the market distribution and marginal movements of income among individuals do not affect welfare.

welfare maximized with a cost re-allocation such that $\Delta u(y_i^0) = \overline{\Delta u}$ for all households. In terms of income, this implies a specific notion of fair burden given by

$$r_i = r(y_i^0) = \frac{\overline{\Delta u}}{u'(y_i^0)} = \frac{\overline{u'\Delta y}}{u'(y_i^0)} \tag{4}$$

As shown in the appendix, we can use these values of a_i in (3) and r_i in (4) to produce $\bar{u}'W_0$ and $\bar{u}'W_1$ in (1) and (2) through a bit of manipulation and parameter assumptions. Hence, Slesnick provides one basis for choosing our welfare measures.

How does this expression for fair burden, $(\bar{u}'/u_i')\overline{\Delta y}$, vary across households with different income levels? That depends on the shape of the utility function. Suppose we assume iso-elastic utility, where

$$u_i = u(y_i) = (1-\tau)^{-1} y_i^{1-\tau}, \tag{5}$$

so $u'(y_i) = y_i^{-\tau}$. Consider two households, rich (R) and poor (P) where $y_R^0 > y_P^0$. Given the above expression above for r_i, we have $r(y_R^0)/r(y_P^0) = (u_R'/u_P')^{-1} = (y_R^0/y_P^0)^{\tau}$. When $\tau = 1$ (i.e., log utility), the welfare-maximizing cost allocation is an equal percentage of income for all households. When $\tau > 1$, the rich household should pay a disproportionate share of income than the poorest. That is, $r(y_R^0)/r(y_P^0) > y_R/y_P$. When $\tau < 1$, the rich household still pays more in absolute terms but less than a proportionate share of income relative to the poorest.

The Negishi weights have another important and related consequence for the Slesnick welfare function. Imagine we are examining an outcome where $0 > \Delta u_i(y_i) - \overline{\Delta u} > \Delta u_j(y_j) - \overline{\Delta u}$. Both households are faring worse than the average burden, $\overline{\Delta u}$. But household j is bearing a more extreme adverse burden. Consider a small transfer of income to household i from j. Along the lines of the Pigou-Dalton principle, we would want this transfer to improve welfare, since it would reduce the more extreme deviation from the average utility change without affecting individuals other than *i* and *j*. Based on the Negishi weights, this will be true so long as $\rho > 0$. That is, the derivative of the second term in (2) for a reallocation dy from *i* to *j* would be

$$(1+\rho)\left(a_j\left|\Delta u_j - \overline{\Delta u}\right|^\rho u'\left(y_j^0\right) - a_i|\Delta u_i - \overline{\Delta u}|^\rho u'(y_i^0)\right)dy$$

$$= \frac{(1+\rho)}{\sum_i u'\left(y_i^0\right)^{-1}}\left(\left|\Delta u_j - \overline{\Delta u}\right|^\rho - |\Delta u_i - \overline{\Delta u}|^\rho\right)dy,$$

which is positive so long as $\rho > 0$, given the larger deviation in utility for household *j*. If $\rho = 0$, Pigou-Dalton holds only weakly. Welfare is not improved by such transfers, but neither is it reduced. In that case, we do not care about more extreme burdens.

This point highlights the importance of ρ in the Slesnick function. The form $(\sum_i a_i|\Delta u_i - \overline{\Delta u}|^{1+\rho})^{\frac{1}{1+\rho}}$ is an example of a power mean. This simplifies to an arithmetic mean of when $\rho = 0$ and standard deviation when $\rho = 1$, our two formulations of interest. More generally, the expression converges to the maximum value of $|\Delta u_i - \overline{\Delta u}|$ as $\rho \to \infty$.[7] In other words, ρ governs the degree of aversion to extremes of inequality in the Pigou-Dalton sense, versus a general aversion to differences, however small or large. Larger values of ρ will imply more concerns about extreme deviations, while $\rho = 0$ only cares about the average deviation.

The only remaining parameter is γ. A value of $\gamma \geq 0$ simply reflects the relative importance of equity, measured by the second term, and overall cost, measured by the first. If γ is zero, there is no concern for the distribution of costs. For large values of γ, we are increasingly willing to accept a higher overall cost to society in order to achieve a more equitable burden. Slesnick picks γ to be as large as possible while still satisfying the criterion that a Pareto

[7] See Chapter 3 of Bullen (2003)

improving policy raises welfare regardless of the distribution. As we show in the appendix, this amounts to $\gamma = 1$ for W_0. We require $\gamma \leq 1$ to be consistent with the Pareto criterion but are otherwise agnostic.

Prospect theory

The welfare measure W_0 in (1) can also be motivated by prospect theory. Kahneman and Tversky (1979) argue that gains or losses are evaluated relative to a reference point and welfare exhibits loss aversion and diminishing sensitivity. Consistent with prospect theory, Tversky and Kahneman (1992) offer a value function for a gain or loss *x* with the power function form $v(x) = x^\alpha$ for $x \geq 0$, and $v(x) = -(1+\lambda)(-x)^\beta$ for $x < 0$, where $\alpha > 0, \beta > 0,$ and $\lambda > 0$ implies loss aversion.[8]

Let us create an aggregate welfare function W_{PT} reflecting the principles of prospect theory, with underlying assumptions analogous to those in W_S. Assume that $\alpha \approx \beta \approx 1$.[9] Furthermore, the gain or loss is assessed relative to an individual reference point, r_i, so *x* in the value function is given by $x = \Delta y_i - r_i$ where Δy_i is again the income change for household *i*. We write an aggregate welfare function, including individual reference points and loss aversion:

[8] al-Nowaihi et al (2008) show that preference homogeneity in the presence of loss aversion then requires $\alpha = \beta$. Diminishing sensitivity would require $\alpha \in (0,1)$, implying risk aversion over gains and risk seeking over losses.

[9] This assumption implies that marginal utility is locally flat, allowing for straightforward aggregation.

$$
\begin{aligned}
W_{PT} &= \underbrace{N^{-1}\sum_{i=1}^{N} r_i}_{\substack{\text{reference}\\ \text{points}}} - (1+\lambda)\underbrace{\left(N^{-1}\sum_{i=1}^{i^*}\left|\Delta y_i - r_i\right|\right)}_{\text{aggregate losses}} + \underbrace{\left(N^{-1}\sum_{i=i^*+1}^{N}\left|\Delta y_i - r_i\right|\right)}_{\text{aggregate gains}} \\
&\equiv \overline{\Delta y} - \lambda\left(N^{-1}\sum_{i=1}^{i^*}\left|r_i - \Delta y_i\right|\right)
\end{aligned}
\qquad (6)
$$

where we have ordered individuals from greatest loss to greatest gain, i^* is the last individual suffering a loss (e.g., $\Delta y_i < r_i$ for $i \leq i^*$ and $\Delta y_i \geq r_i$ for $i > i^*$), and $\lambda > 0$ for loss aversion. In order to simplify to the second line of (6), suppose the reference point is some notion of a fair cost burden of a particular aggregate cost, $\sum \Delta y_i$, so $\sum r_i = \sum \Delta y_i$ as before. With that assumption, the sum of the absolute value of losses equals the sum of the absolute value of gains: i.e., $\sum_{i=1}^{i^*}(r_i - \Delta y_i) = \sum_{i=i^*+1}^{N}(\Delta y_i - r_i)$. We can further rewrite expression (6) to show that a mean-preserving increase in the absolute deviations of outcomes reduces welfare:

$$
\begin{aligned}
W_{PT} &= \overline{\Delta y} - \lambda\left(N^{-1}\sum_{i=1}^{i^*}\left|r_i - \Delta y_i\right|\right) + \frac{\lambda}{2}\left(N^{-1}\sum_{i=1}^{i^*}\left|r_i - \Delta y_i\right| - N^{-1}\sum_{i=i^*+1}^{N}\left|\Delta y_i - r_i\right|\right) \\
&= \overline{\Delta y} - \frac{\lambda}{2}\left(N^{-1}\sum_{i=1}^{N}\left|r_i - \Delta y_i\right|\right)
\end{aligned}
$$

Replacing $\gamma = \lambda/2$, this is the same expression as W_0 in (1). Prospect theory leads to a more generic notion of fair burden, r_i, which is otherwise determined by equal utility change in the Slesnick formulation. On the other hand, the Slesnick framework allows more easily the incorporation of a more general notion of inequality aversion that can be sensitive to more extreme deviations from the welfare-maximizing burden and motivates the alternative W_1 in (2).

Discussion

Ultimately, using either the Slesnick or prospect theory approach requires assigning values to what are at best subjective parameters of the social welfare function. These subjective parameters include the degree of inequality aversion γ in (3) or loss aversion λ in (6), the notion of utility curvature τ in (5) or fair burden r_i in (6), and the aversion to extreme inequality ρ in (3).

One approach is to assume values for some parameters in order to provide relatively simple expressions, as we have done for ρ in (1) and (2) and partly for τ in (2). For others, such as fair burden r_i and inequality aversion γ, we leave them unspecified for the moment. We then return to discuss these parameters as we present numerical welfare results and compare policies.

A somewhat different approach is to use this discussion to recognize that the distribution of Δy_i by decile is generally what matters for welfare. We can then present this information graphically and using summary statistics for various policy alternatives. The end user applies their own judgement and values to draw conclusions, rather than trying to choose parameters.

We now turn to a policy application to highlight these approaches.

Modeling household outcomes under different electricity sector policies

To make our observations about horizontal equity applied to energy regulation concrete, we consider a stylized example of alternative policies designed to achieve the same carbon emission outcome in the electric power sector: cap-and-trade (CAT) and tradable performance standards (TPS). This choice of policies is a particularly relevant question for stakeholders. Both types of policies have been proposed for the electric power sector over the past decade (Waxman 2009; Bingaman 2012). The Clean Power Plan also provided states with options for both rate-based and mass-based trading—in other words, tradable performance standards or cap-and-trade. China is currently implementing a tradable performance standard in the power sector, even as other countries have embraced cap-and- trade (Pizer and Zhang 2018).

To construct our example, we first present a simple analytic model to highlight different household outcomes under the two policies and to relate those outcomes to a small number of parameters. We then use data from the Consumer Expenditure Survey and other sources to quantify the household outcomes. Subsequently, we show how these effects look when viewed through the lens of the welfare functions developed in the previous section.

Simple electricity sector model

Our economic framework for comparing policies is a partial equilibrium model of the power sector. On the demand side, we focus on the case of perfectly inelastic electricity demand by each household. It may seem strange to abstract from notion of demand response, which eliminates any aggregate cost advantage of CAT, the Pigouvian policy, over TPS in our simple model. That is, the underlying point of the paper is that there is an equity-efficiency trade-off, and here we assume there is no efficiency advantage of CAT.

However, a necessary condition for an equity-efficiency trade-off is that equity effects are large enough that TPS could be preferred. By focusing on the case of inelastic demand, we focus on just how large the equity concern might be. Most importantly, fixing electricity demand simplifies our exposition. Each household's loss of real income equals its individual increase in electricity costs minus its share of any allowance value rebated directly to households.

On the supply side, we assume constant-returns-to-scale (CRS) technology with unit cost determined by the carbon price. This allows us to capture the key features that concern us. On the one hand, we want there to be an increase in the cost of electricity associated with a carbon price under either TPS or CAT. On the other hand, we want to capture different electricity price effects when the associated allowance value is either rebated in the electricity price under TPS or assigned to households under CAT. These are the salient features of more complex models we aspire to emulate, such as Burtraw and Palmer (2008).

Formally, let p_z be the electricity price, and p_m be the allowance price. Let C_0 be unit production costs in the absence of regulation. Market-based regulation adds two components: unit abatement costs (*UAC*) and unit emissions payments (*UEP*). If *TAC* is total abatement costs, then $UAC = TAC / Z$, where *Z* is the (fixed) aggregate generation. Assuming cost minimization over a CRS technology, marginal abatement costs (*MAC*) are equal to the price ($\partial TAC / \partial M = p_m$). That is, we treat pollution like any other input that has to be purchased at price p_m and assume other input prices are fixed.

Total emissions payments are $TEP = p_m M$, where *M* is total emissions after responding to the regulation. Similar to converting *TAC to UAC,* unit emissions payments are defined as $UEP = TEP / Z = p_m (M / Z)$. That is, *UEP* equal the emissions price multiplied by the average emissions intensity per unit of generation. Since even freely allocated allowances have an opportunity cost, this component of the unit cost increase occurs regardless of how permits are allocated, and whether they arise under TPS or CAT. We refer to the emission payments interchangeably as emission rents or allowance value.

We see these cost components in Figure 1, where M_0 is the emission level when $p_m = 0$. As the electricity sector begins to pay a positive price $p_m > 0$ for their emissions, *M*, producers will begin to reduce emissions by $M_0 - M$. This incurs an abatement cost, the area under the *MAC* schedule, highlighted by region *TAC* in the figure.

Electricity producers also face a cost p_m for emissions that occur, M, highlighted by region *TEP*. We have drawn the figure for total generation, so we must scale the increase in total production costs by $1/Z$, the fixed total electricity demand, to relate to the change in unit costs of electricity production.

Notably, for all but very deep reduction targets, the size of the emissions rents is much larger than the total abatement costs (*TEP* >> *TAC*). Thus, market-based policies create the potential for large redistributions, based on the allocation of these rents.

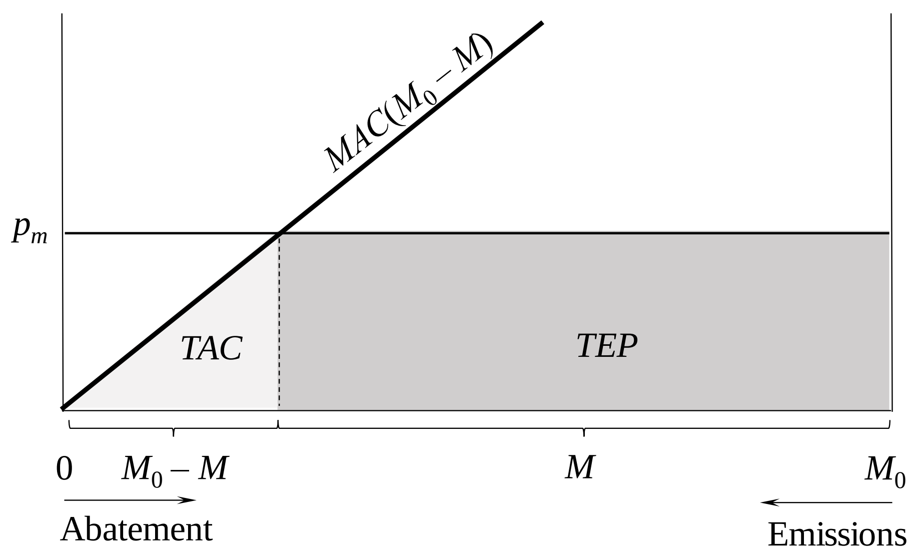


Figure 1. Comparing emissions rents, *TEP,* with compliance costs, *TAC*.

With CAT, the increase in electricity prices due to the regulation equals the sum of the *UAC* and *UEC*:

$$\Delta p_z^{\text{CAT}} = UAC + UEP$$

where the superscript CAT reflects the outcome under cap-and-trade. Allowance values are allocated in lump-sum fashion, so their distribution does not affect behavior or electricity prices. Let us assume that the total allowance value *TEP* is rebated to each household *i* based on an assigned share s_i. That is, each household receives $s_i TEP$.

A TPS sets a performance rate *R* (expressed in pollution / unit of electricity). Each unit of generation is allocated allowances equal to this benchmark, *R*, and through trading an equilibrium is reached where the overall average emissions intensity equals the performance rate, or $M / Z = R$. Fixing total emissions to be the same under both policies, emission prices and total emissions payments are as before (p_m and *TEP*). However, under TPS this allowance value is rebated as a subsidy to electricity production. At the unit level, this subsidy equals $p_m M/Z = p_m R = UEP$. This subsidy is passed on to consumers and serves to mitigate the electricity price increase:

$$\begin{aligned}\Delta p_z^{\text{TPS}} &= UAC + UEP - p_m R \\ &= UAC\end{aligned}$$

Here, the superscript TPS reflects the outcome under the tradable performance standard. That is, the unit cost increase is driven only by the abatement cost, not the allowance rent.

With this supply model, we can now turn to household outcomes. With households $i \in \{1, \ldots, N\}$, let Z_i represent household *i*'s fixed electricity consumption such that $Z = \sum_i Z_i$. As noted above, fixing this consumption implies that a household's loss of real income equals the increased cost of electricity minus any share of the allowance allocation. Under TPS this is given by $\Delta y_i^{\text{TPS}} = -UAC \cdot Z_i$.

Under a CAT, the real income change is given by $\Delta y_i^{\mathrm{CAT}} = -\left(UAC + UEP\right) Z_i + s_i \cdot TEP$; that is, the added cost of buying the fixed electricity demand Z_i subtracted from the household's share of allowance value.

The difference between the policy outcomes for household *i* thus depends on whether the value of the household's share of the allowance revenues exceeds its share of electricity consumption: $\Delta y_i^{\mathrm{CAT}} - \Delta y_i^{\mathrm{TPS}} = (s_i \cdot TEP - Z_i \cdot UEP) = (s_i - Z_i / Z) TEP.$ On net, both shares sum to one, so $\sum_i (s_i - Z_i/Z) = 0$ and aggregate costs for both polices are given by $N\overline{\Delta y} = \sum_i \Delta y_i = -TAC$.

At this point, for expositional purposes, we fix $s_i = 1/N$; that is, equal per-household rebates. This cap and dividend approach is consistently suggested in various carbon pricing schemes (Inglis 2009; Larson 2015; Blumenauer 2017; Baker III et al. 2017). Nonetheless, we return to this assumption at the very end of our analysis.

Table 1. Hypothetical policies for numerical analysis

	Effect on household *i*
Tradable performance standard (TPS)	$\Delta y_i^{TPS} = -(TAC / N)(Z_i / \bar{Z})$
Cap and trade (CAT): with per household rebate	$\Delta y_i^{CAT} = -(TAC / N)(Z_i / \bar{Z}) + (TEP / N)(1 - Z_i / \bar{Z})$
Difference (CAT minus TPS)	$\Delta y_i^{CAT} - \Delta y_i^{TPS} = (TEP / N)(1 - Z_i / \bar{Z})$

Note: *TAC* is total abatement cost, *TEP* is total emission payment, Z_i is household *i*'s electricity expenditure, $\bar{Z}$ is average electricity expenditure, and *N* is the total number of households.

Table 1 summarizes the household outcomes for each policy and how they relate to summary cost parameters and data, with $\bar{Z} = Z / N$ as mean electricity consumption. Based on our model and assumptions, the TPS distributes the abatement costs according to electricity consumption shares, and the CAT policy adds a net emissions rent that is positive for households with below-average electricity consumption. These are the key differences that we want to capture. We now turn to the data that will allow us to quantify our earlier analytic results. In particular, we need to approximate the distribution of Z_i/Z, which equals the household share of total electricity expenditures in the population, and to choose the cost parameters *TAC* and *TEP*.

Household data and mitigation cost

To provide a basis for likely variation in consumption of electricity and other demographics necessary for the calculations in Table 1 and further discussion, we use US consumer expenditure data. In particular, we turn to the 2014 Consumer Expenditure Survey, or CEX (BLS 2014). This is a rolling, quarterly survey, where a representative sample of US

Table 2: Summary statistics for numerical exercise

	observations	Mean	Std. dev.	Min	Max
Electricity ($, C_0Z_i)	1,086	1,037	844	0	5,907
Log(Electricity)	1,036	6.72	0.764	3.64	8.68
Expenditures ($, y_i^0)	1,086	35,936	32,518	1902	330,237
Log(Expenditures)	1,086	10.2	0.821	7.55	12.71
Electricity share (%, C_0Z_i/y_i^0)	1,086	3.97	3.59	0	28.1

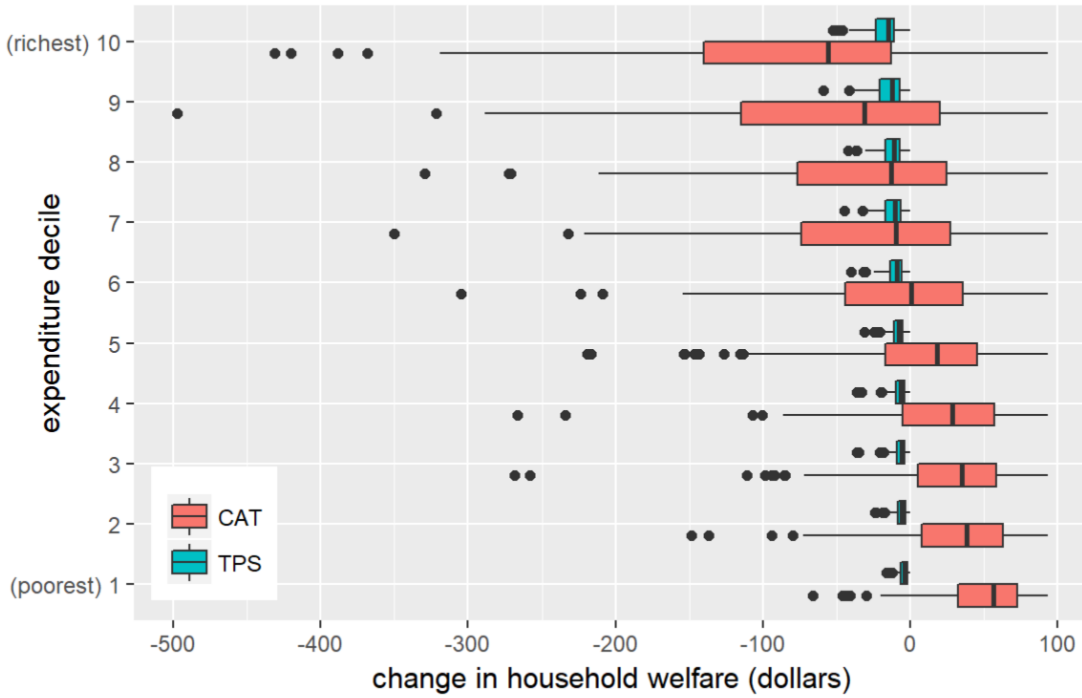
households enters each quarter and remains in the survey for five quarters. We compute the total expenditure on electricity and total expenditures overall for the calendar year. We include only survey respondents who participated for the entire year (1,086). That is, we first match household respondents on their household identifier for each quarter of 2014 and keep only those households observed for all four quarters. We sum reported expenditures on electricity over these four quarters, as well as total expenditures.[10] Table 2 summarizes the data. We also indicate the notation we have been using that corresponds to each viable.

From Table 1, we also need to specify the mitigation costs and rents, *TAC* and *TEP*. Based on recent analysis (EIA 2009), a reasonable assumption is that cap-and-trade regulation on carbon dioxide might raise electricity prices on the order of 10 percent. Based on other analysis (Burtraw and Palmer 2008), a reasonable assumption is that the actual cost (without the allowance revenue) is perhaps 10 percent of that (i.e., a 1 percent increase in electricity prices). Thus we choose *TAC* = 0.01 times the electricity expenditure in the sample and *TEP* = 0.09 times total electricity expenditure. Given the summary statistics, where the mean electricity expenditure ($C_0\bar{Z}$) was roughly \$1,000 per household, we have *TAC/N* = \$10 and *TEP/N* = \$93. With these data and parameters in hand, we now turn to our results.

Policy comparison and welfare measures

We plug the CEX data on Z_i and parameters *TAC/N* and *TEP/N*, all just discussed, into the expressions in Table 1 for income effects by household. This yields distributions for Δy_i^{CAT} and Δy_i^{TPS} across households. Figure 2 presents these distributions graphically by decile using

[10] Total expenditures (TOTEXPPQ) include all outlays by households for goods and services as well as contributions to pensions.



Note: See Table 1. *TAC/N* = \$10, *TEP/N* = \$93, and the distribution of $Z_i/\bar{Z}$ is as described in Table 2. Boxes indicate interquartile range (IQR, 25th to 75th percentile). Vertical lines in the middle of the boxes indicate median. Horizontal lines, or whiskers, show range of values outside the IQR, up to 1.5x the IQR. Dots indicate each individual values beyond whiskers. For normally distributed data, such dots should have a frequency of ~1%.

Figure 2. Comparison of cap-and-trade (CAT) and tradable performance standard (TPS), in dollars

box-and-whisker plots where CAT is red and TPS is blue and deciles are arranged from poorest decile at the bottom to richest at the top. Two observations stand out. First, while the TPS outcomes are all negative (consistent with Table 1), the CAT outcomes tend to be positive for poorer households. Because poorer households have smaller electricity expenditures, the per capita rebate under CAT leads to these positive welfare effects for the majority of households in the lower half of the income distribution. Second, the range of outcomes is much larger within each decile under CAT than TPS. For example, some households in even the poorest decile see negative effects under CAT. Among the poorest four deciles roughly one-quarter remain worse off.

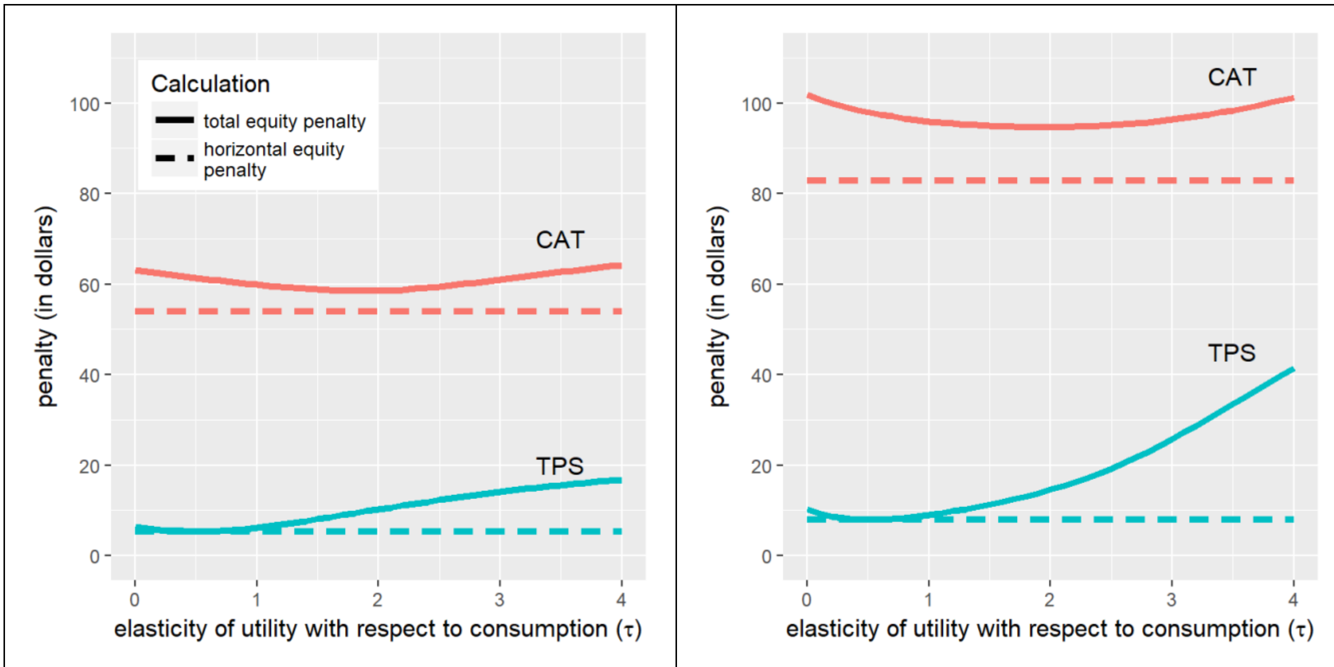
Note that most alternatives to a per capita rebate would vary the rebate by income. Such policies would shift the box plots for each expenditure decile but not change the spread within the decile—given the spread is unrelated to income. We return to the idea of alternative ways to define the rebate later, but arguably none of these alternatives would fundamentally change the distinction that CAT creates more within-decile variability than TPS.

This observation reflects one of the key, practical points of this paper. While CAT policies can generally achieve any desired cost distribution across income groups, including positive outcomes, on average for the poorest, they cannot avoid significant variation. The range of outcomes is inherently much larger under CAT than TPS because the rents *TEP* tend to be much larger than the mitigation costs *TAC*, and because there is significant within-income-group variation in household electricity use. Once rents enter electricity prices, this large, within-income-group variation will be difficult to ameliorate.

Measuring welfare

How might this variability translate into welfare considerations? We now turn to our operational welfare measures, focusing mainly on the equity penalty term (recalling the first welfare term equals *TAC/N* = –$10 for both policies). The equity penalty arises from the failure of the actual distribution of household burden Δy_i to match the notion of fair burden r_i.

Based on the welfare measures in (1) and (2), we first calculate a "total equity penalty." We focus on Slesnick's definition (4) of fair burden $r_i = (\overline{u'}/u_i')\overline{\Delta y}$. Using $u'(y_i) = y_i^{-\tau}$ as in (5) leads to $r_i = \left(y_i^0 / \bar{y}\right)^{\tau} \overline{\Delta y}$. As noted above, fair burden will rise as a share of income at higher income levels when $\tau > 1$ (and the reverse when $\tau < 1$).



Equity Penalty $N^{-1}\sum\lvert\Delta y_i - r_i\rvert$ in W_0 in (1)	Equity penalty $\sqrt{N^{-1}\sum_i(\bar{y}/y_i^0)(\Delta y_i - r_i)^2}$ in W_1 in (2)

Note: For the solid line (total equity penalty), $r_i = \left(y_i^0/\bar{y}\right)^{\tau}\overline{\Delta y}$. For the dashed line (horizontal equity), $r_i = \overline{\Delta y}_{d(i)}$.

Figure 3. Effect of varying power utility parameter τ on equity penalty.

We noted that one of the useful features of our welfare definitions (1) and (2) is that they allow a decomposition into horizontal and vertical equity effects. In addition to the total equity penalty, we also compute the "HE penalty" that arises when we substitute a reference point equal to the average burden in each household's income decile $r_i = \overline{\Delta y}_{d(i)}$, where $d(i)$ identifies that decile (e.g., d maps individuals $\{1,\ldots, N\}$ into deciles $\{1,\ldots,10\}$).

Figure 3 first looks at how the equity penalty varies with fair burden as defined by τ. The solid lines in the figure show the total equity penalty and dashed lines show the HE penalty. Values of the both penalties appear along the vertical axis for CAT (red) and TPS (blue) policies with values of τ indicated along the horizontal axis. The left panel shows W_0 and the right panel W_1. Note that the HE penalty does not vary with τ, having replaced the expression of r_i that depends on τ.

We make three observations. First, the penalties are uniformly larger for W_1 (right panel) than W_0 (left panel) As noted earlier, these two welfare measures can be derived as specific cases ($\rho = 0$ and $\rho = 1$) of the Slesnick welfare function (3). The parameter ρ determines defines the extent to which the penalty tends to the average absolute deviation versus the more extreme absolute deviations, with higher values of ρ putting more weight on more extreme values. Thus. it should not be surprising that, by assuming a larger ρ, W_1 yields larger equity penalties.

Second, the total equity penalty varies with τ reaching a minimum in both the left and right panels at $\tau \sim 0.5$ for the TPS and $\tau \sim 2$ for the CAT. The reflects the idea that there is a value of τ where the fair burden over initial income levels, determined by τ, most closely matches each policy's actual distribution of average outcomes. At that value of τ the penalty is minimized.

Finally, the CAT penalty is much higher than the TPS penalty. The welfare difference, ~\$50 per household for W_0 and ~\$80 for W_1, is large compared to the average cost (ignoring equity) of \$10 given by *TAC/N*. This difference is clearly driven by the HE component, as the dashed lines account for most of the difference between CAT and TPS.

This is the other key, practical point of this paper. The variation of policy impacts within income groups can be large under Pigouvian pricing. That variation translates into larger, negative horizontal equity impacts for Pigouvian versus non-Pigouvian policies. Depending on the weight given to these impacts (e.g., the parameter γ in (1) and (2)), it appears large enough to overwhelm some differences in efficiency (e.g., differences in $\overline{\Delta y}$).

Despite its usefulness, missing from Figure 3 is an understanding of how different deciles contribute to the equity penalty. Is the penalty, particularly the horizontal component, more sensitive to variation in the rich or poor? We answer that question in Table 3. We present the same information as in Figure 3 broken down by decile, for a single value of $\tau = 1$ (log utility). That is, eight columns (5-8 and 11-14) in Table 3 take the values reported by the 8 lines in Figure 3 for $\tau = 1$, reproduces them in the last row, and then breaks them down by income decile in the remainder of the table.

The breakdown by deciles shows that poorer deciles contribute more to the equity penalty under W_1 versus W_0. Looking at each policy-total equity and policy-HE combination, the penalties uniformly decline moving from W_0 to W_1 for the richest decile but increase for the poorest. E.g., horizontal equity for CAT declines from \$83 to \$59 for the richest decile and increases from \$27 to \$84 for the poorest. This may seem counter-intuitive. The range of dollar values is actually largest for the richest decile in Figure 2. Indeed, looking at the standard deviation of burden by decile, reported in columns 4 and 10, we see the largest values for the richest deciles. As discussed previously, the equity penalty in W_1 is generally larger than the penalty in W_0 because it puts more weight on extreme deviations and the richest decile has the largest deviations.

Table 3. Summary by decile of total and horizontal equity penalties ($) from cap-and-trade (CAT) versus tradable performance standard (TPS), using W_0 and W_1

income decile	avg. income $\overline{y^0_{d(i)}}$	fair burden defined by log utility	CAT						TPS					
			avg. burden $\overline{\Delta y}^{CAT}_{d(i)}$	standard deviation	penalty*				avg. burden $\overline{\Delta y}^{TPS}_{d(i)}$	standard deviation	penalty*			
					Total equity (W_0)	*HE* (W_0)	Total equity (W_1)	*HE* (W_1)			Total equity (W_0)	*HE* (W_0)	Total equity (W_1)	*HE* (W_1)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
10	108,100	–31.2	–84.5	107.2	84.5	83.5	65.6	59.3	-17.8	10.7	15.4	8.4	10.3	5.9
9	60,900	–17.6	–59.4	102.0	78.9	80.6	84.3	78.5	-15.3	10.2	8.6	8.1	8.1	7.9
8	46,100	–13.3	–33.5	79.3	61.8	62.4	71.7	69.5	-12.7	7.9	6.3	6.2	7.0	7.0
7	36,600	–10.6	–32.7	82.0	62.5	64.2	84.3	81.3	-12.6	8.2	6.3	6.4	8.4	8.1
6	29,300	–8.5	–13.5	69.1	51.3	52.3	77.2	77.0	-10.7	6.9	5.1	5.2	8.1	7.7
5	23,500	–6.8	2.3	67.3	52.7	49.5	84.3	83.6	-9.1	6.7	4.7	5.0	8.8	8.4
4	19,100	–5.5	16.6	59.7	51.6	42.7	87.3	81.9	-7.7	6.0	4.3	4.3	8.7	8.2
3	15,300	–4.4	21.7	61.3	52.9	42.1	103.7	95.5	-7.2	6.1	4.2	4.2	10.4	9.6
2	11,300	–3.3	30.8	47.1	49.8	35.6	104.0	83.7	-6.3	4.7	3.9	3.6	9.9	8.4
1	6,400	–1.9	47.9	34.2	54.3	27.1	161.1	84.5	-4.5	3.4	3.2	2.7	10.2	8.5
Total	35,900	–10.4	–10.4	74.1	60.1	53.9	96.2	79.9	-10.4	7.4	6.2	5.4	9.1	8.0

*Penalties computed using penalty term in equations (1) for W_0 and (2) for W_1. Total equity (columns 5, 7, 11, 13) define $r_i = (y_i^0/\bar{y})\overline{\Delta y}$ (summarized by decile in column 4). Horizontal equity (*HE*, columns 6, 8, 12, 14) define $r_i = \overline{\Delta y}_{d(i)}$ as given in column 3 (for CAT) and 9 (for TPS).

The explanation lies in the weight $(\bar{y}/y_i^0)$ appearing in (2). In our derivation of (2) from the Slesnick welfare function we embed an assumption of log utility with our assumption of ρ = 1 (see appendix). This implies a concern about variation in changes as a *share* of incomes. This leads us to down-weight rich households (where shares have a larger denominator) and up-weight poorer ones (where shares have a smaller denominator) based on $(\bar{y}/y_i^0)$. Usefully, the difference between the standard deviation by decile reported in columns 4 and 10 versus the exact HE calculation in columns 9 and 14 is largely a factor of $(\bar{y}/\overline{y_{d(\iota)}})$.

A more general approach to horizontal equity

We have discussed our results in terms of the welfare measures W_0 and W_1. However, stakeholders may understandably be hesitant to embrace the ethical judgements of economists embedded in W_0 and W_1. This includes the choice of ρ in (3) and τ in (5), as well as the general "black box" nature of the calculations.

Conveniently, all of the information necessary to make welfare judgements is contained in Figure 2 and columns 3-4, 6, 9-10, and 12 in Table 3. Both provide information about outcomes by decile, including the central tendency and measures of spread, for each policy. Stakeholders can decide for themselves how much to weight deviations within deciles (horizontal equity) as well as how to value the central tendency of each decile (vertical equity) versus some objective. They need not buy into our particular assumptions embedded in W_0 and W_1. At the same time, the information is arguably consistent with the use of a welfare approach.

We view this approach as similar to the use of Lorenz curves. Lorenz curves represent simple a simple summary of income inequality as well as defining a specific welfare measure. However, stakeholders can use Lorenz curves to understand inequality within society, and to make policy choices among alternatives, without necesssarily using the particular welfare measure or adopting its particular ranking of outcomes.

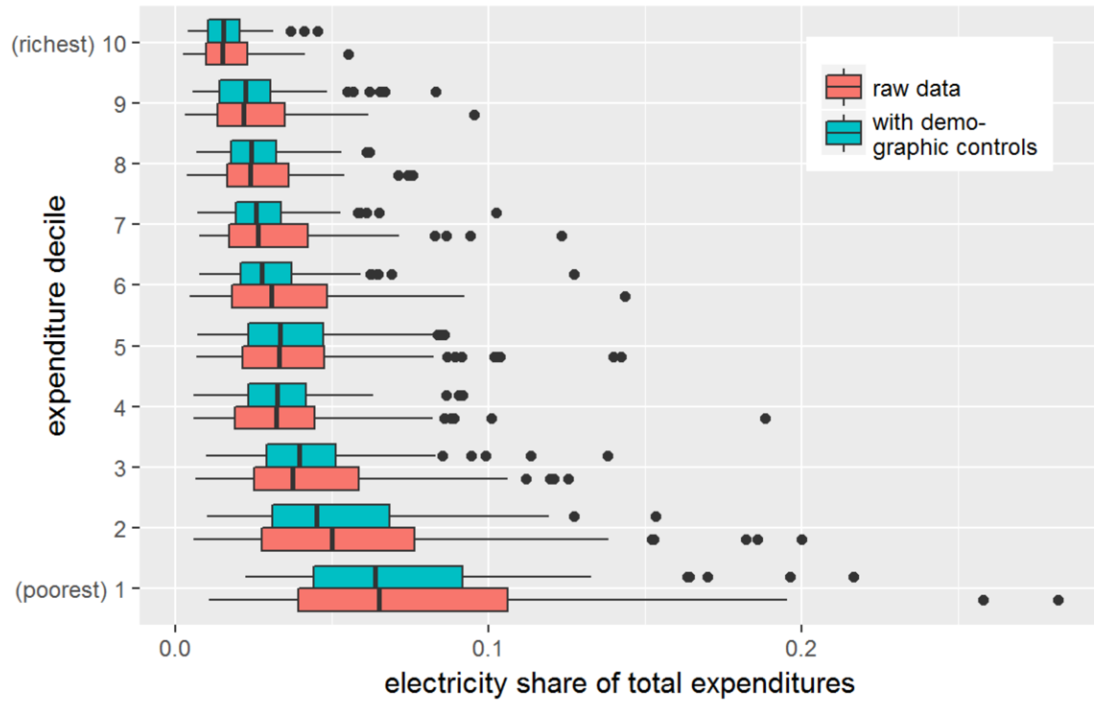
Could CAT do better with more targeted rebates

In our stylized comparison of CAT and TPS, we have suggested that CAT can have a higher equity penalty when we consider horizontal inequality. This stems from the fact that there is considerable within-income-group variation in electricity use, and CATs higher impact on the electricity price. Lurking in this result is an assumption that we cannot fix these unequal effects after the fact, with targeted rebates that address this heterogeneity.

But could we? To what extent might we improve income-based redistribution and move towards targeted, within-income-group rebates? We know that electricity expenditures vary with household size and location, among other observed variables. How well can we predict electricity expenditures, controlling for income?

We explored this question by taking our data from the CEX and trying to predict logged electricity expenditures. More precisely, we took all of the household characteristics contained in the CEX interview survey, converting categorical variables to indicators, and replacing missing geographic identifiers with zeros.[11] This resulted in a set of 133 variables. With this enhanced data set, we had 879 complete observations (of 1,036 original observations). We then used the LASSO algorithm with cross-validation to choose the best predictive model that is robust to concerns about multiple hypothesis tests (James et al. 2013). We found 35 variables, including

[11] This includes all variables listed as "Consumer Unit (CU) Characteristics" in the data dictionary. For many observations, geographic identifiers are omitted to protect confidentiality in the public-use data sets. For our purposes, available identifiers (e.g., 0/1 variables for particular states or PSUs) can be useful predictors and missing values simply become a reference group where we do not know the location.



Note: Boxes indicate interquartile range (IQR, 25th to 75th percentile). Vertical lines in the middle of the boxes indicate median. Horizontal lines, or whiskers, show range of values outside the IQR, up to 1.5x the IQR. Dots indicate each individual values beyond whiskers. For normally distributed data, such dots should have a frequency of ~1%.

Figure 4. Box-and-whisker plots of electricity share by decile, highlighting the predictability of electricity expenditures using household income and other covariates (red is actual data; blue uses machine-learning algorithm to choose covariate controls)

total expenditures, useful in predicting electricity use.[12] Most of these are geographic of family composition indicators. However, all of these variables together predict about half of total variation in electricity use (R-squared of 0.56), leaving considerable residual variation.[13]

We present these results graphically in Figure 4. As in Figure 2, we use box-and-whisker plots to show the variability within expenditure deciles but here the horizontal axis expenditure

[12] This includes 19 geographic identifiers (2 regional indicators, 9 state indicators, and 8 PSU indicators), 6 income variables (log expenditures, 2 rank variables, INC_RANK and ERANKHM, and 3 income category indicators), 4 family size and age variables (family size, 1 indicator for all children >17 years, 2 family type indicators), and 6 variables describing the housing location (2 population size indicators, a rural-urban indicator, and 2 indicators of housing tenure).

[13] Total expenditures alone predict 28%. These other 35 variables roughly double the predictive power.

share rather than dollar expenditures. Red boxes are the raw data and blue boxes are the residual variation after using these 35 variables to predict electricity expenditures. Visually, considerable variation remains (incrementally, income alone explains 28% of the variation).

While suggestive, this is still not a complete picture. The government will have considerable more information about individuals. For example, they may have more finely tuned geographic identifiers.[14] Such data may allow more precise targeting of rebates. However, we suspect considerable variation will remain, based on variation in housing age, design, etc., unless one is willing to turn to historic electricity use. And at some point, an increasingly complex scheme may become impractical.

Conclusion

Our principal motivation has been to highlight that Pigouvian policies in the energy sector may have large and often overlooked distributional consequences. In particular, they tend to raise energy prices and lead to greater variation in household-level effects *within* income groups. These consequences are difficult to remedy through typical redistribution schemes. Other policies to reduce pollution can have smaller effects on energy prices, and hence smaller distributional consequences, even as they have higher aggregate costs to society.

Should this variation in household costs within income groups matter? Traditional welfare notions tend to focus on overall costs to society. Distributional effects matter to the extent that they change the underlying income distribution and make it more or less equitable. That is, transfers from rich to poor are welfare-improving given any level of overall costs. In this

[14] We observe state or primary sampling unit for 88% of our sample. Nonetheless, many states (e.g., California) have a wide range of climate zones.

paper, we have highlighted the notion of fair burden as an alternative to traditional welfare notions. Fair burden emphasizes how the cost of a public good should be shared across households, treating those with similar income (and other characteristics) similarly, *without* an implicit welfare reward for redistribution from rich to poor. Generally, we would expect a regulation to entail a non-negative burden for all households. As fair burden regards even-handed changes, the approach does place special emphasis on the pre-regulation distribution of income as the primary basis for assigning burden. We have shown how to operationalize this approach by positing welfare measures based on Slesnick (1989) and prospect theory (Kahneman and Tversky 1979). Both theories lead to a penalty based on how household income changes deviate from fair burden or another reference point, incorporating both horizontal and vertical equity.

We made these ideas concrete through the stylized comparison of two policy options that have been proposed to address carbon dioxide emissions in the electricity sector—cap-and-trade with equal per-household rebates (CAT) and tradable performance standards (TPS). CAT indeed leads to much wider variation in income changes across all income groups. Applying our welfare measure, we found that the associated CAT equity penalty is several times that of the TPS and potentially larger than efficiency advantages (about which we only speculate). A lingering question is whether more targeted rebates under CAT could ameliorate the otherwise large variation in income changes that underlies the penalty. Turning to available data, the answer appears to be no.

Our welfare measure does not tell us how much to weight the equity penalty versus concerns about efficiency, a question of ethical and societal preferences. Such measures can also appear to be a “black box” to stakeholders, making them unappealing. For these reasons, we also emphasize practical and intuitive ways to present the relevant data that drives our welfare measures, including tables and figures describing the distribution of outcomes by decile. This approach is analogous to the use of Lorenz curves to describe the income inequality associated with traditional welfare notions. By making the relevant outcomes easy to understand, stakeholders can draw their own conclusions directly, and largely consistent with our welfare measures.

Given the oft-apparent disconnect between economists promoting Pigouvian policies and policymakers choosing non-Pigouvian alternatives, this paper raises an interesting possibility. Perhaps horizontal equity and distributional effects are something that policymakers have recognized for some time, and that only economic analysis has tended to overlook.

References

Adler, Matthew D. 2013. "The Pigou-Dalton Principle and the Structure of Distributive Justice." Working Paper. Duke University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2263536.

Auerbach, Alan J., and Kevin A. Hassett. 2002. "A New Measure of Horizontal Equity." *American Economic Review* 92 (4): 1116–25. https://doi.org/10.1257/00028280260344650.

Baker III, James A., Martin Feldstein, Ted Halstead, N. Gregory Mankiw, Henry M. Paulson Jr, George P. Shultz, Thomas Stephenson, and Rob Walton. 2017. "The Conservative Case for Carbon Dividends." The Climate Leadership Council. https://www.clcouncil.org/.

Bingaman, Jeff. 2012. *S.2146 - 112th Congress (2011-2012): Clean Energy Standard Act of 2012*. https://www.congress.gov/bill/112th-congress/senate-bill/2146.

Blumenauer, Earl. 2017. *H.R.3420 - 115th Congress (2017-2018): American Opportunity Carbon Fee Act of 2017*. https://www.congress.gov/bill/115th-congress/house-bill/3420/text.

Bullen, P. S. 2003. *Handbook of Means and Their Inequalities*. 2nd ed. Mathematics and Its Applications. Springer Netherlands. //www.springer.com/gp/book/9781402015229.

Bureau of Labor Statistics (BLS). 2014. "Consumer Expenditure Survey." 2014. https://www.bls.gov/cex/.

Burtraw, Dallas, David A. Evans, Alan Krupnick, Karen Palmer, and Russell Toth. 2005. "Economics of Pollution Trading for SO2 and NOx." *Annual Review of Environment and Resources* 30 (1): 253–89. https://doi.org/10.1146/annurev.energy.30.081804.121028.

Burtraw, Dallas, and Karen Palmer. 2008. "Compensation Rules for Climate Policy in the Electricity Sector." *Journal of Policy Analysis and Management* 27 (4): 819–47. https://doi.org/10.1002/pam.20378.

Cohen, Richard E. 1995. *By Richard E. Cohen - Washington at Work: Back Rooms and Clean Air: 2nd (Second) Edition*. Longman.

Elkins, David. 2006. "Horizontal Equity as a Principle of Tax Theory." *Yale Law & Policy Review* 24 (1): 43–90.

Energy Information Administration (EIA). 2009. "Energy Market and Economic Impacts of H.R. 2454, the American Clean Energy and Security Act of 2009." https://www.eia.gov/analysis/requests/2009/hr2454/.

Inglis, Bob. 2009. *Raise Wages, Cut Carbon Act of 2009*. https://www.congress.gov/bill/111th-congress/house-bill/2380/text.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning: With Applications in R*. 1st ed. 2013, Corr. 7th printing 2017 edition. New York: Springer.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–291.

Kaplow, Louis. 1989. "Horizontal Equity: Measures In Search Of A Principle." *National Tax Journal* 42 (2): 139–54.

———. 1992. "Note on Horizontal Equity, A." *Fla. Tax Rev.* 1: 191.

———. 2000. "Horizontal Equity: New Measures, Unclear Principles." NBER Working Paper 7649. National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/7649.html.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *The Quarterly Journal of Economics* 121 (4): 1133–1165.

———. 2007. "Reference-Dependent Risk Attitudes." *The American Economic Review* 97 (4): 1047–1073.

———. 2009. "Reference-Dependent Consumption Plans." *American Economic Review* 99 (3): 909–36. https://doi.org/10.1257/aer.99.3.909.

Larson, John. 2015. *H.R.3104 - 114th Congress (2015-2016): America's Energy Security Trust Fund Act of 2015*. https://www.congress.gov/bill/114th-congress/house-bill/3104/text.

Michaelson, Zachary. 2015. "Biases in Choices about Fairness: Psychology and Economic Inequality." *Judgment and Decision Making* 10 (2): 198.

Mill, John Stuart. 1871. *Utilitarianism*. Longmans, Green, Reader, and Dyer.

Mueller, Dennis C. 2003. *Public Choice III*. 3 edition. Cambridge ; New York: Cambridge University Press.

Negishi, Takashi. 1960. "Welfare Economics and Existence of an Equilibrium for a Competitive Economy." *Metroeconomica* 12 (2–3): 92–97. https://doi.org/10.1111/j.1467-999X.1960.tb00275.x.

Nowaihi, Ali al-, Ian Bradley, and Sanjit Dhami. 2008. "A Note on the Utility Function under Prospect Theory." *Economics Letters* 99 (2): 337–39. https://doi.org/10.1016/j.econlet.2007.08.004.

Pigou, A. C. 1928. *A Study in Public Finance*. Macmillian.

Pizer, William A., and Xiliang Zhang. 2018. "China's New National Carbon Market." *AEA Papers and Proceedings* 108: 463–67. https://doi.org/10.1257/pandp.20181029.

Poterba, James M. 1991. "Is the Gasoline Tax Regressive?" *Tax Policy and the Economy* 5: 145–64.

Rausch, Sebastian, Gilbert E. Metcalf, and John M. Reilly. 2011. "Distributional Impacts of Carbon Pricing: A General Equilibrium Approach with Micro-Data for Households." *Energy Economics*, Supplemental Issue: Fourth Atlantic Workshop in Energy and Environmental Economics, 33, Supplement 1 (December): S20–33. https://doi.org/10.1016/j.eneco.2011.07.023.

Sen, Amartya. 1979. "Utilitarianism and Welfarism." *The Journal of Philosophy* 76 (9): 463–89. https://doi.org/10.2307/2025934.

Slesnick, Daniel T. 1989. "The Measurement of Horizontal Inequality." *The Review of Economics and Statistics* 71 (3): 481–90. https://doi.org/10.2307/1926905.

Thaler, Richard. 1980. “Toward a Positive Theory of Consumer Choice.” *Journal of Economic Behavior & Organization* 1 (1): 39–60.

US Government. 1993. “Executive Order 12866. Regulatory Planning and Review.” US Government.

Waxman, Henry. 2009. *H.R.2454 - 111th Congress (2009-2010): American Clean Energy and Security Act of 2009*. https://www.congress.gov/bill/111th-congress/house-bill/2454/text.

Weinzierl, Matthew C. 2012. “Why Do We Redistribute so Much but Tag so Little? The Principle of Equal Sacrifice and Optimal Taxation.” NBER Working Paper 18045. National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/18045.html.

Westort, Peter J., and Janet M. Wagner. 2002. “Toward a Better Measure of Horizontal Equity.” *Journal of the American Taxation Association* 24 (1): 17.

Appendix: Simplifying Slesnick to produce our welfare measures

With our choice of a_i to be Negishi weights, the result that $\overline{\Delta u} = \overline{u'}\overline{\Delta y}$, and the local approximation $\Delta u_i = u_i' \Delta y_i$, we can rewrite Equation (3) as

$$W_s = \overline{u'}\overline{\Delta y} - \gamma \left(N^{-1} \sum_i \left(\frac{u_i'}{\overline{u'}} \right)^{-1} |u_i' \Delta y_i - \overline{u'}\overline{\Delta y}|^{1+\rho} \right)^{\frac{1}{1+\rho}}$$

Rearranging slightly:

$$W_s = \overline{u'} \left(\overline{\Delta y} - \gamma \left(N^{-1} \sum_i \left(\frac{u_i'}{\overline{u'}} \right)^{\rho} |\Delta y_i - (u_i')^{-1} \overline{u'}\overline{\Delta y}|^{1+\rho} \right)^{\frac{1}{1+\rho}} \right)$$

Defining $r_i = (u_i')^{-1}\overline{u'}\overline{\Delta y}$, for $\rho = 0$ we then have

$$W_s = \overline{u'} \left(\overline{\Delta y} - \gamma N^{-1} \sum_i |\Delta y_i - r_i| \right) = \overline{u'} W_0$$

Further, assuming $u(y_i^0) = \ln(y_i^0)$ and $r_i = (u_i')^{-1}\overline{u'}\overline{\Delta y} = (y_i^0/\bar{y})\overline{\Delta y}$, for $\rho = 1$ we then have

$$W_s = \overline{u'} \left(\overline{\Delta y} - \gamma \sqrt{N^{-1} \sum_i \frac{\bar{y}}{y_i^0} (\Delta y_i - r_i)^2} \right) = \overline{u'} W_1$$

We note that

$$\frac{dW_0}{d\Delta y_i} = \frac{1}{N} - \gamma \frac{1}{N} sign(\Delta y_i - r_i)$$

This will be non-negative so long as $\gamma \leq 1$.