August 2016

Equity versus Efficiency in Energy Regulation

Carolyn Fischer

William A. Pizer

Duke University

Duke University Energy Initiative

Box 90467

Durham, NC 27708

www.energy.duke.edu

Equity Effects in Energy Regulation

Carolyn Fischer and William. A. Pizer *

Abstract

Choices in energy regulation often involve options with divergent impacts on household energy prices, particularly regarding electricity. At the same time, household energy use varies widely, across income and demographics, as well as within any easily identifiable group. This suggests that these choices can have significant distributional effects to be weighed alongside efficiency concerns. In this paper, we explore the equity consequences of changing energy prices and various options for accounting for these consequences.

Key Words:

JEL Classification Numbers:

* Carolyn Fischer is… William A. Pizer is Professor, 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.

1. Why equity effects matter

In the acid rain trading program, an agreement was made to cap SO_2 emissions from coal-fired power plants at a specific level. Following this agreement, and a general notion of how to allocate emission rights, attention shifted to horse-trading among the companies to address the exact distribution of burden <citation>. The acid rain trading program has been viewed as a significant success and motivated significant attention to market based policies in the decades that followed, particularly surrounding carbon dioxide. This includes literally dozens of programs around the world <cite WB report>.

One important difference between the acid rain trading program and CO2 programs is the effect of electricity and the distribution of costs. In the acid rain program, the price of electricity was largely unaffected and the horse-trading companies were the ones who felt the effects of the regulation and allocation choices. CO2 programs, however, have the potential to raise electricity prices significantly. Households feel the direct impact of regulation as well as the choices about allowance or revenue allocations. The horse-traders are elected officials, not those feeling the effects of the regulation or allocation.

Under these circumstances, it seems hard to ignore equity and distributional concerns. In the acid rain program, it may have been sufficient to focus on efficiency, because equity concerns were addressed directly by the affected parties—Coasian bargaining at its best. But when elected or appointed representatives are making these decisions, they need to decide who pays. Moreover, the politics of achieving any outcome may hinge such decisions being acceptable to key stakeholders.

Part of the dilemma is that Pigouvian pricing policies, absent transfers, can involve redistributions that are many times the net gain or loss to society. Burtraw and Palmer (2008) find that a pricing policy has social costs of roughly $0.5 billion annually, while consumers and producers lose more than $21 billion in payments to the government. Among firms, the aggregate loss is $3 billion, but one group of facilities gains $6 billion while another group loses $9 billion. In contrast, performance standards involve smaller redistributions but lead to higher social costs. If distributional effects matter and are not addressed, the efficiency advantage of pricing policies is less compelling. Moreover, beliefs about individual and net social benefits could be lower still. Even if policymakers intend to address distributional concerns to avoid inequity, lack of confidence in the outcomes means public support for the program could falter.

This paper will consider various ways to evaluate both the overall cost of a policy and the inequality of its burdens. Importantly, we will consider social welfare functions that allow distribution of costs, not just final outcomes, to matter. Moreover, we want to understand how concern about unequal burdens might translate into toleration of higher overall costs.

Ultimately, welfare cost (like cost-benefit analysis itself) is not a singularly decisive metric for policymaking. Unless it is particularly intuitive, it may also obscure underlying information that would be relevant. For that reason, we also consider what kind of information might be presented in order to inform better decision making.

Our paper ultimately highlights four points. First, different forms of energy regulation will have different effects on energy prices. Pricing pollution, in particular, will typically have larger price impacts than other forms of regulation. Second, efforts to address unequal impacts are limited by variation that is not linked to easily observed, or easily targeted, information. Third, social welfare can be reduced by policies with unequal burdens across households. These types of social welfare measures point to a tradeoff between the higher efficiency but lower equity of pollution pricing. Finally, we speculate that the trade-offs could be more severe if either (a) stakeholders are skeptical about whether or how revenue from pollution pricing will occur, and (b) the creation of larger losers under pricing policies leads to more vocal opposition. Taken together, these points emphasize that without considering distributional consequences, economic analyses of various policy alternatives risk excluding important effects on overall welfare.

2. Different policy designs: Tradable permits versus tradable performance standards

A motivating example for thinking about equity concerns in energy regulation is the idea of emissions pricing versus performance standards for regulating energy-related emissions. Both entail direct compliance costs, but emissions pricing, like taxes or tradable permits, also imposes charges on the remaining emissions. Performance standards forego these charges on embodied emissions. Tradable performance standards have the advantage of equalizing emissions prices across regulated entities, while essentially forgiving the costs associated with an average rate of emissions (the performance standard). The implicit rebate is passed on to consumers in the form of lower costs (or smaller cost changes), whereas tradable permit systems pass on the opportunity cost of emissions. Energy-related pollution lends itself particularly well to tradable performance standards because the energy output is often easy to measure, so regulations targeting "tons per megawatt-hour" are sensible. By focusing on tradable permits and tradable performance standards we compare two policies that both feature marginal cost equalization. However, these two policies, particularly applied to energy regulation, highlight an important feature that concerns us: regulations that create significant rents that directly influence households and those that do not.

In particular, consider a simple economy with two goods, E and X. Think of the goods as energy and everything else, respectively. For simplicity, we assume a population of N households that are endowed with different amounts of X indicated by X_n^0. Moreover, each household has different preferences over E and X summarized by $U_n(E, X)$. Energy is produced using both X and emissions M through a constant returns to scale technology defined by the unit cost function $C(p_x, p_m)$. Without loss of generality, we let X be the numeraire, $C(1, p_m) = C(p_m)$, and $v_n(p_e, w_n)$ is the indirect utility function, where w_n is wealth.

This formulation captures the key features we want to emphasize—and very little more. In particular, households differ from one another both in terms of wealth (e.g., X_n^0) and, even conditioning on wealth, their consumption of energy versus other goods. There is only one factor of production, so there are no source-side distribution effects outside the initial endowment.

Let X be the numeraire. Consider an initial equilibrium given by $p_m = 0$.

$$p_e = C(0)$$

$$u_n = v_n(p_e, x_n^0)$$

Now consider a cap on pollution M with shadow price λ. The unit price of energy equals the unit cost plus the shadow value of the embodied emissions, M/E per unit:

$$p_e^C = C(\lambda) + \lambda M / E$$

Household n's weath includes a share s_n of the allowance revenues, leading to utility

$$u_n = v_n(p_e^C, x_n^0 + s_n \lambda M)$$

Now consider a rate-based regulation, with $M/E < R$, or $M \leq R \times E$, with the shadow value μ. In this case, the unit price of energy equals the average unit cost of meeting the regulation $C(\mu)$, without the implicit carbon charge:

$$p_e^R = C(\mu)$$

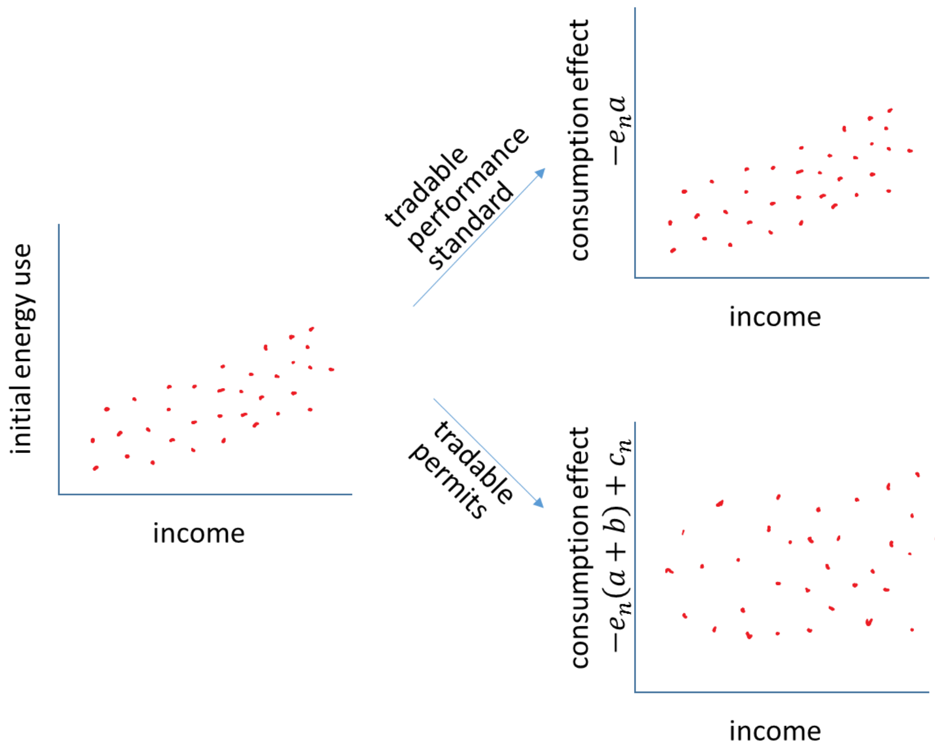
$$u_n = v_n(p_e^R, x_n^0)$$

Assuming they achieve the same emissions level, $\mu \geq \lambda$ and $p_e^R < p_e^C$: since pricing the emissions embodied in energy encourages conservation and substitution away from the energy good, emission reductions can be achieved with less effort to reduce emissions intensity than with the performance standard (see, e.g., Fischer and Fox 2007). The first inequality will be strict unless demand for E is completely inelastic. That is, if there is no demand response, the credit price in the tradable performance standard just rises to achieve the same abatement within the energy production process. Meanwhile, the energy price increase must be lower under the tradable performance standard.

What does all of this math mean? A tradable performance standard will require a higher credit price to achieve the same emissions level because it does not raise the price of energy as much—unless energy demand is perfectly inelastic.

Let us consider the case of perfectly inelastic energy demand. Let e_n be household n's share of aggregate energy demand. Note the change in aggregate consumption will be $E\big(C(\lambda) - C(0)\big)$ in both cases. Household n's change in consumption of X, however, will vary. Under tradable permits, it is given by $s_n \lambda M - e_n E\big(C(\lambda) + \lambda M/E - C(0)\big)$. Under a tradable performance standard, it is given by $-e_n E\big(C(\lambda) - C(0)\big)$. Hence the difference for household n is given by $(s_n - e_n)\lambda M$, which depends on whether the value of its share of the allowance revenues exceeds its payments for embodied emissions. Moreover, $\sum_n (s_n - e_n) = 0$. That is, for tradable permits, there is a higher cost per unit of energy due to the carbon price, but the emissions payments are offset by rebates. The net effect is zero for the population as a whole, but can be positive or negative for any particular household. In contrast, the tradable performance standard ensures that revenues are rebated directly per unit.

The figure below highlights this idea. The left side posits a particular distribution of energy use across household (vertical axis) that varies with income (horizontal axis). The right side shows an imagined effect of alternative tradable performance standards and tradable permits. In the former case (upper right), there is a particular price increase in energy, translated into cost now represented on the vertical axis. In the latter case (lower right), there is a *larger* increase in energy prices—and a corresponding larger vertical spread—coupled with a particular choice of rebate. In this case, we imagined trying to equalize effects across income.



This emphasizes two important questions that arise when thinking about the *distributional* differences between tradable permits and tradable performance standards: How large is the potential redistribution of rents, b, (where $b \equiv \lambda M$) compared to the direct cost increase, a, (where $a \equiv E\left(C(\lambda)-C(0)\right)$), and how might $(s_n - e_n)$ vary across households? That is, how much higher is the energy price increase caused by the rents compared to the aggregate consumption costs, and how might those rents be redistributed?

One can quickly get an idea about the first question by examining analysis of tradable permit systems applied to the energy sector, comparing the magnitude of permit value to changes in aggregate consumption. Of course, in the real world, there is heterogeneity in production, potentially non-market behavior, and source-side income effects. Not all the rents will flow through energy prices. But with those caveats in mind, the ratio is still quite large <cite>.

The second question is more difficult. Most real-life examples of proposed reforms <cite> suggest s_n might be a per capita refund, or be connected to income through a cut to income or social security taxes (as suggested in the figure). That does not mean other possibilities are not possible, but it does suggest a natural assumption would be that s_n be tied to income.

3. Welfare and decision making

A key observation is that energy policies in general, and tradable permits and performance standards in particular, will have varying impacts both across and within groupings of households based on income. Economic theory tells us that tradable permits well generally have lower overall costs than tradable performance standards, but it does not speak to the distribution of these varying impacts. Tradable permits are often heralded because they can flexibly redistribute rents through other tax changes and/or direct transfers, seeming to negate distributional concerns. However, it is unlikely they can address all variability of concern in practice. The variability within income groups seems particularly difficult to resolve, requiring household allocations based on geography, historical energy use, or some other proxy. And, while creating these flexibly distributable rents tradable permits exacerbate variability.

Do we care about this? Is it not enough that tradable permits have lower overall costs? Should we worry that variability might be higher under tradable permits, perhaps offsetting their advantage in overall costs? There are two ways to answer this question: based on ethical considerations or based on political considerations. We begin with ethical considerations and social welfare theory, asking two particular questions: "What does this particular theory imply about the welfare-maximizing cost distribution?" And, "how do particular deviations from this welfare-maximizing distribution—particularly within and across income groups—lead to lower welfare?"

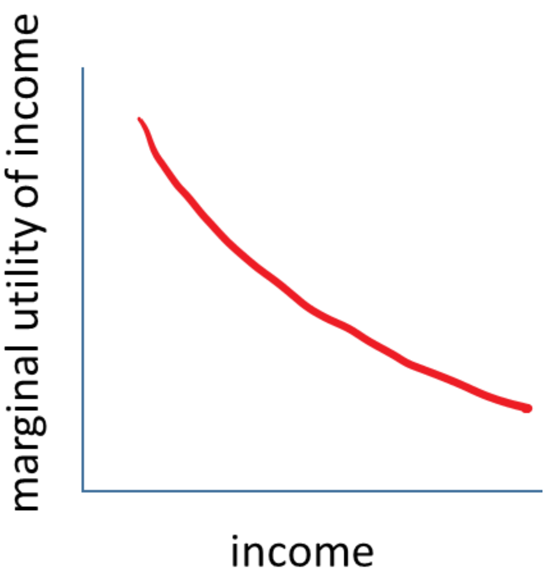
The long history of welfare economics has tended to emphasize a conceptual framework for defining how individual outcomes, and preferences over those outcomes, are mapped into social preferences.[1] A fairly typical approach would be to define a mapping from individual utility u_i to social welfare $= W(u_1, u_2, \ldots, u_N)$ where the u_i are cardinal and interpersonally comparable utility measures. Social welfare theory has then gone on to consider various forms of the function $W(.)$.

An ethical way to answer the question of "do we care" is to ask whether $W\left(u_1', u_2^i, \ldots u_N'\right) > W(u_1'', u_2'', \ldots u_N'')$ where u_i' are the outcomes under tradable permits and u_i'' are the outcomes under tradable performance standards. Assuming that what matters is lower cost overall, and not distributional concerns, the condition will be true. If distributional concerns do matter, or are not addressed by whatever approach is taken with rents, the condition can be

[1] See Chapter 23 of Mueller (2003) for a discussion of traditioanl social welfare theory.

false. The question would seem to come down to both choosing how to manage rents under tradable permits and the form of $W(.)$.

Peeling back the onion a bit more, however, this may not be quite right. Imagine for a moment a simple utilitarian welfare function $W = \sum_i u_i$. Moreover, assume that marginal utility for each household is given by $(x_i)^{-\tau}$, the amount of non-energy consumption of household *i* raised to a negative power, where $\tau \geq 0$. Looking at the figure below, it is clear that welfare changes are going to be dominated by any transfers from better-off households to worse-off households. That is, cost-preserving policies that include transfers from richer to poorer households will unambiguously raise welfare until incomes are equalized. The implication is that, to the extent we care about welfare impacts of energy policies other than total cost (e.g., $\tau > 0$), those effects will tend to be dominated by redistribution from rich to poor. Is that the correct notion for energy policy? In particular, welfare could be higher for an energy policy that had higher total costs but achieved more redistribution.



Consider an even simpler example. Suppose a policy called "even" costs household A and B $10 each while policy called "skew" costs household A $20 and household B $0. Both have average costs of $10, but which has higher welfare? The above discussion suggests that it depends on A and B's incomes. So long as A is at least $10 richer than B, we know that the marginal utility of A is strictly lower than the marginal utility of B as one moves from policy "even" to policy "skew", making "skew" the preferred policy. But does that really make sense?

Consider a more extreme policy called "robinhood" that costs household A $110 and household B $–100. Suppose A has $101,000 in income and B has $100,000. The marginal utility of B is strictly higher over the range of costs, so this will raise welfare even more than the other two policies. Do most people think this would be fair? Should we be using energy policy to redistribute income? Moreover, suppose A has higher income but a higher cost of living?

Geographic cost of living adjustments, as well as other household differences, is conceptually addressed by using household equivalence scales (Jorgenson and Slesnick 1984). In practice, it is not clear how effective that will be. Even fixing those concerns, differences in income are still determining the inequality of an energy policy.

There is an alternative ethical framework. The traditional line of thought regarding welfare presumes inequality is determined by the distribution of total income, which may be better or worse under alternative policies. One alternative is that the status quo represents an important ethical benchmark, and that what matters for welfare and equity is deviations from the status quo. This idea, often associated with Nozick's (1974) entitlement theory of justice, is not without controversy. In particular, it is difficult to defend when the status quo is in some sense "unfair."[2] At the same time, if we view the current tax system as society's choice of how to address income inequality, it may be appropriate to consider the status quo distribution absent energy policy as the correct benchmark for measuring inequality. Alternatively, one might have in mind to consider an equitable distribution of costs without rewarding redistribution (or negative costs).

This idea of a benchmark distribution is closely related to the idea of horizontal equity. Horizontal equity is the idea that it is ethically desirable to treat equally situated households equally. Implicit in this idea is that the initial situation has some significance, and simply swapping household positions without changing the overall distribution should lower welfare. Many of same criticisms follow.

As a matter of implementation, at least two social welfare functions have been used in the literature. Slesnick (1989) uses a welfare function of the form

$$W_S = \underbrace{\frac{1}{N}\sum_i \Delta u_i}_{\text{Average impact}} - \underbrace{\gamma\left(\frac{1}{N}\sum_i \left|\Delta u_i - \overline{\Delta u}\right|^{1+\rho}\right)^{\frac{1}{1+\rho}}}_{\text{Loss from inequality}} \tag{1}$$

where Δu_i are the changes in utility for each individual from a reference level or status quo, N is the number of households, ρ is a horizontal inequality aversion parameter, and γ is a weight.[3]

The first term in (1) represents a standard utilitarian welfare function that simply gives the average impact, with equal weights for each household. In this term, any increase in individual utility implies higher social welfare, and any concern about inequality stems from the curvature of individual utility functions (vertical inequality).

The second term introduces the notion of inequality from the standpoint of the reference case or horizontal inequality. To the extent that changes in individual utility are different from the mean change in individual utility, welfare is penalized. In other words, with this welfare

---

[2] See for example Kaplow (1989)

[3] We have adjusted his formulation for consistency with the assumption that $\rho > 1$.

measure, if only one individual experiences higher utility, the first term rises but the second term falls.

The parameter $\rho > 0$ captures concern about inequality and the more extreme values of $|\Delta u_i - \Delta \bar{u}|$ across households. For example, $\rho = 1$ would be akin to subtracting off the standard deviation of the individual utility changes. As $\rho \to \infty$, the welfare measure emphasizes the most extreme deviation.[4] At $\rho = 0,$ the second term implies subtracting the average absolute deviation.

The parameter $\gamma > 0$ serves as a drag on the redistribution of utility. It is required so that not only are more equitable distributions of changes in individual utility higher in social welfare, but also that a dollar transfer from a household with an above average *change* in utility to a household with a below average change will increase total welfare. For example, suppose we choose u_i to simply be income, y_i, (e.g., ignoring traditional "vertical" inequality), so

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

In this case, any coefficient in front of the second term would satisfy this property. But more generally it requires a particular (at least minimum) choice of γ. For example, we might consider u to be the log of income:

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

In both cases, there will be a particular allocation of costs that maximizes welfare. In the first case, it is when every household has exactly the same dollar change in income. In the second, it is when every household has exactly the same *percent* change in income. Here, you can also see the role of the weights γ. Shifting a dollar from the richest to the poorest would, according to the first term, raise W by $(y_{min}^{-1} - y_{max}^{-1})/N$, by raising the average marginal utility. For inequality concerns to dominate, the weight γ needs to be large enough so that the second term counteracts that increase and welfare decreases, reflecting the welfare loss of an unequal burden. This condition precludes a smooth transition from a welfare function that values inequality in this way (with the second term) to one that does not (without the second term). As noted, the parameter ρ captures *how* inequality in changes is valued, in terms of all deviations being equally valued versus large deviation mattering more. But variation always matters.

---

[4] This indicator relates to Rawlsian utility, which emphasizes the utility of the worst off; however, here the largest deviation—which could be positive or negative and be attributed to a household with high or low reference utility—is what gets the greatest weight.

A different approach is taken in Auerbach and Hassett (2002). They focus on welfare in terms of income and how to modify a "traditional" welfare function to account for horizontal inequality. If individuals were grouped into I bins with N_i individuals, where each of the N_i's has an identical income, y_i^0, in a reference case, their welfare function for any policy case is given by[5]

$$W_{AH} = \left[\sum_i N_i \left(\frac{1}{N_i} \sum_j y_{ij}^{1-h} \right)^{\frac{1-v}{1-h}} \right]^{\frac{1}{1-v}} \tag{2}$$

This framework blends the issue of horizontal equity (valued by the parameter h) with that of vertical equity (valued by the parameter v). Note that the inside term is in essence the adjusted income for group i, multiplied by a horizontal equity measure reflecting "local" inequity around i. In the reference case $y_{ij}^0 = \bar{y}_i^0$ and everyone in the group has the same income. In the policy case, there is some average change income $\bar{y}_i - \bar{y}_i^0$. But welfare in the new state is adversely affected by the dispersion of the y_{ij} based on the curvature of the function given by γ.

One implication of this model is what an "equitable" distribution of costs looks like. It clearly lowers welfare to introduce any variation within those groups. Moreover, the welfare maximizing distribution of costs would (a) put all the burden on the wealthiest group, and (b) encourage transfers from rich to poor as much as possible. Thus, like the Slesnick model, it penalizes "horizontal inequity" within groups. However, it maintains a traditional view of equity across groups, emphasizing redistribution unless $v = 0$. Another implication is that there is a smooth transition to the standard welfare model. As $\gamma \rightarrow \rho$, this becomes a standard utilitarian welfare function.

In reality, every household has a different income, so this first approach is not practical without modification. Auerbach and Hassett propose to instead consider a window around each observation, and how much increased inequality there is in that window relative to the reference case. We review this case in the Appendix.

For the present analysis, if we simplify to look just at horizontal equity, with $v = 0$, we have

[5] Again, we adjust some variable names to avoid confusion with others already designated.

$$W_{AH}{}' = \sum_i N_i \left(\frac{1}{N_i} \sum_j y_{ij}^{1-h} \right)^{\frac{1}{1-h}}$$

$$= \sum_i N_i \bar{y}_i \underbrace{\left(\frac{1}{N_i} \sum_j \left(\frac{y_{ij}}{\bar{y}_i} \right)^{1-h} \right)^{\frac{1}{1-h}}}_{\equiv H_i}$$

That is, welfare is just income, times an adjustment factor (H_i) depending on a weighted average of the horizontal equity measures.

One issue with this approach for examining energy policy is that it is hard to interpret. Generally, energy policy is going to have a relatively small impact on income. The variation may be a few percent, or fractions of a percent, of income. In that case, $y_{ij}/\bar{y}_i$ will be close to 1. Indeed, even for the examples in Auerbach and Hassett focused on income taxes, all of their estimates of H_i are above 99%. In their discussion, focused on comparing horizontal inequality across income percentiles or over time, that may be okay. It is less clear how one would present this alongside of other measures of vertical inequality or overall costs.

Even when $v = 0$, within a small neighborhood, it will raise welfare to redistribute (at least a little) from the rich to the poor so long as $h > 0$. In their own examples, Auerbach and Hassett use a window of around +/–10% to define horizontal equity around any particular observation. With that window, we will see below that it is possible to begin seeing this effect.

As a final welfare-based example, we consider a variation on Slesnick. Rather than writing the initial term as the average change in utility, suppose we wrote it as a change in "aggregate utility" for an "aggregate representative consumer". That is,

$$W_{FP} = \underbrace{\Delta u}_{\substack{\text{Change for}\\\text{average}\\\text{household}}} - \underbrace{\gamma \left(\frac{1}{N} \sum_i \left| \Delta u_i - \overline{\Delta u} \right|^{1+\rho} \right)^{\frac{1}{1+\rho}}}_{\text{Loss from inequality}} \tag{3}$$

where u is just the individual utility function applied to average per capita income. Unlike the Slesnick expression, where the first term depended on the distribution of costs across individuals (unless $u_i = y_i$), the first term now only depends on aggregate costs. Then the second term introduces a penalty based on the variation in utility changes across individuals. That is, this implies that welfare is always maximized when $\Delta u_i = \Delta \bar{u}$ for all individuals, for a given total cost, regardless of the choice of ρ or γ. This allows us to both choose how we regard inequality (in a Rawlsian or utilitarian way, applied to changes) and how we weight inequality, generally, versus overall costs, without a constraint on γ having a minimum value.

We can summarize these four different representations of welfare in the table below, highlighting some of the key differences.

Table 1: Summary of welfare functions for equity analysis

	Equity notion	*Shifting $ from rich to poor always raise welfare?*	*Penalize unequal treatment of similar income?*	*Smooth transition to concern over HE?*
Standard welfare function	Equalize incomes	Yes	No	NA
Auerbach and Hassett	Equalize incomes; equal treatment of equal initial income	Yes	Yes	Yes
Slesnick	Equalize change in utility	No	Yes	No
This paper	Equalize change in utility	No	Yes	Yes

When one turns to the question of actual decision making, it may be less important to consider explicit social welfare functions or even horizontal inequality measures. Rather, we could ask what information would convey a picture of the distributional effects of a particular policy without being overly burdensome or complex. If we examine how this has been done in the past, we have a couple of examples.

Poterba (1991) is perhaps the first example of an energy tax incidence analysis that tries to get at horizontal equity alongside vertical equity. Like most research <cite vertical equity studies of energy policies), Poterba presents expenditure shares by decile without appealing to a welfare function or summary statistics. Without calling it horizontal equity, he is concerned about the heterogeneity of a gasoline tax within income groups, particularly the lowest. So he adds to his table of expenditures shares by decile (his Table 2) another table highlighting "dispersion" that provides both the fraction of each decile above and below a particular expenditure share (in this case, zero and 0.3; see his Table 3). There, it is clear that the average share of 0.039 for the lowest decile masks that 36% of this decile spend nothing on gasoline while 14% have an expenditure share exceeding 0.1.

Among practitioners, horizontal equity is a more regular concern. The typical approach has been to focus on the coefficient of variation within various income groups (Westort and Wagner 2002). Various approaches to the computing the coefficient of variation have been discussed and are often presented for tax reform proposals.

Taking a more expansive approach, Rausch et al (2011) use graphical figures to present the distributional effects of carbon pricing with various rebate approaches. They present box-and-whisker plots that show both the mean share by decile along with the interquartile range and range of data within 1.5 times the interquartile range on either ride (see our Figure xxx below). In their paper, they use these to highlight that while there is certainly some amount of progressivity and regressivity, with the mean cost by decile ranging from 0 to 0.5% of income, a large number of households experience gains and losses of more than 1%.

In general, such graphical figures may be the best way to present when there are only a few policies to consider. However, we might also consider presenting summary statistics for a few or many percentiles. For example, one could create a table of costs as a share of income by decile, along with summary statistics such as the standard deviation within each decile. Alternatively, one could pick particular deciles of interest—say the 5th, 50th, and 95th. Using a flexible model (e.g., a kernel regression or spline), one could estimate the mean share at those points as well as the standard deviation. We turn to such ideas below.

4. Numerical examples

In this section we consider how different welfare metrics would indicate inequality in the outcomes for hypothetical policies. Following our motivating example, we imagine regulating pollution in the electricity sector based on a very simple model of the economy. Under the assumptions discussed earlier, effects on households can be summarized based on their energy (now electricity) usage and consumption of other non-electricity goods.

To provide a basis for likely variation in household consumption of electricity and non-electricity goods, we make use of US consumer expenditure data. In particular, we turn to the 2014 Consumer Expenditure Survey. We compute the total expenditure on electricity and total expenditures overall for the year. We only include survey respondents who participated for the entire year (1086). The figure and table below summarizes the data.

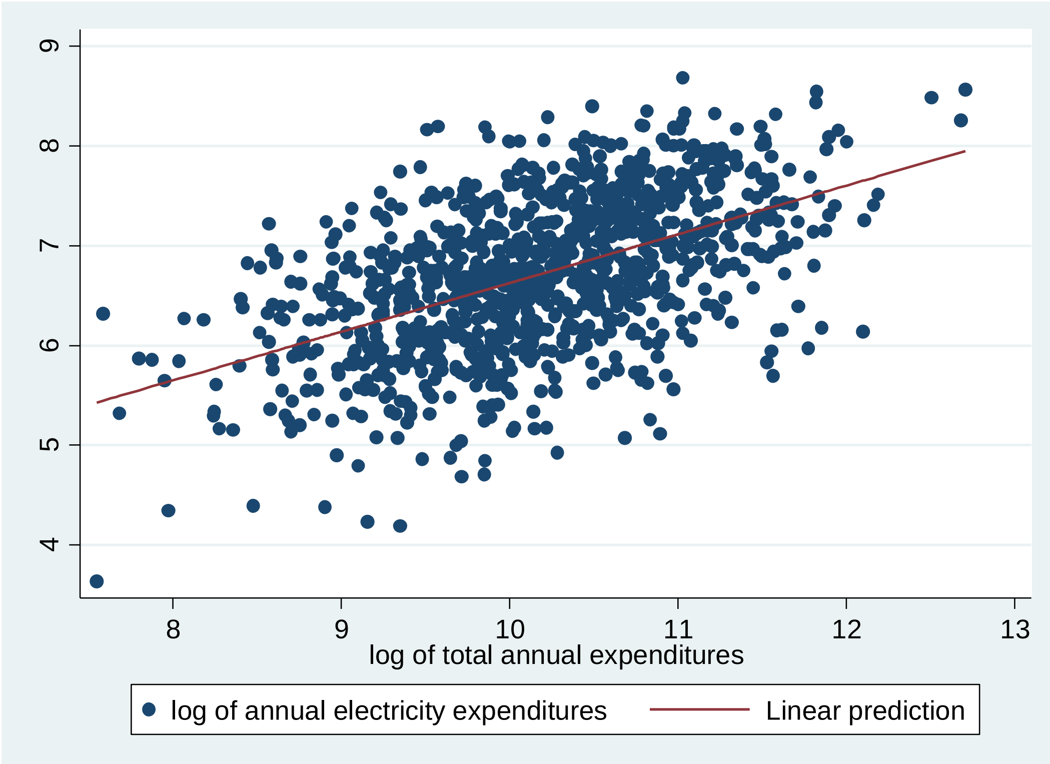


Table 2: Summary statistics for numerical exercise

	observations	mean	Std. dev.	Min	Max
Electricity	1,086	1,037	844	0	5,907
Log(Electricity)	1,036	6.72	0.764	3.64	8.68
Expenditures	1,086	35,936	32,518	1902	330,237
Log(Expenditures)	1,086	10.2	0.821	7.55	12.71

This provides the basic grist for our hypothetical policies. The three policies we have in mind follow from the tradable performance standard and cap and trade described earlier. For simplicity we maintain the assumption that energy demand is completely inelastic. Using the formulas developed in section 2, the tradable performance standard reduces consumption of non-electricity goods by an amount $a \cdot e_i E$. Meanwhile, the cap and trade reduces such consumption by an amount $(a + b) \cdot e_i E$. However, the aggregate revenue $b \cdot E$ is refunded to households in some manner. Our three hypothetical policies are below:

Table 3: Hypothetical policies for numerical analysis

	Impact on household i
Tradable performance standard	$-a \cdot e_i \cdot E$

Cap and trade (per household rebate)	$(-ae_i + b\left(\frac{1}{N} - e_i\right))\ E$
Cap and trade (income based)	$(-ae_i + b\left(f(y_i) - e_i\right))\ E$

Here, we have considered two possible refund mechanisms for the cap and trade. One is to rebate the revenue equally per household. Another is to attempt to equalize costs per household *based on income*. That is, in the above figure, we have indicated a best linear fit of electricity expenditure to income. We use that fitted line to estimate the typical electricity bill for each income level, and rebate the allowance revenue in proportion to that predicted level (e.g., $f(y_i)E$ is the predicted electricity expenditure from the above figure). Given the variation around the line, we know this will under compensate some households and over compensate others.

The final question is to choose the parameters a and b. Based on recent analysis <cite Waxman-Markey etc.> 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 <cite Burtraw or Murray> a reasonable assumption is that the actual costs (without the allowance revenue) is perhaps 10 percent of that (e.g., a 1 percent increase in electricity prices). Thus we choose a = 0.01 times the initial electricity price and b = 0.09 times the initial electricity price.

The figure below shows the effect of the three policies on household non-electricity consumption based on the above assumptions. In particular, the cap and trade with per household rebates is the same as the tradable performance standard, spread out by a factor of ten, and shifted upwards (by the equal per household rebate) to be mean-preserving. The cap and trade with income-based rebates instead shifts higher incomes up more than lower incomes, but is also mean-preserving.

We now examine the various measures of inequality, particularly horizontal inequality described earlier. Based on Auerbach and Hassett, we compute values of H_i as described above and plot them for several choices of w_{ij} (the window for computing local horizontal inequality; see the Appendix) and h (the degree of horizontal inequality aversion). For simplicity, we ignore any aversion to income inequality (v = 0) as discussed above. Thus, absent any horizontal inequality, the welfare costs will generally equal the dollar costs.

Following their example, we let w_{ij} be defined by a normal PDF as a function of the difference between the log of y_i^0 and y_j^0. This implies that a household with $10,000 in income bears similarly on computing H_i at $20,000 as a household with $100,000 income bears on computing H_i at $200,000. The actual choice of standard deviation for window around each observation, as well as the inequality aversion parameter γ are indicated in the table for each of the three policies.

Table 4: Welfare loss calculations with Auerbach and Hassett framework

		Tradable performance standard	*Emissions trading (per capita rebate)*	*Emissions trading (income-based rebate)*
Loss of income (dollars per capita)		10.37	10.37	10.37
Loss of welfare				
bandwidth	*h*			
5%	0.5	10.37	10.39	10.41
	2	10.37	10.44	10.55
	4	10.37	10.51	10.73
10%	0.5	10.37	10.31	10.42
	2	10.38	10.15	10.59
	4	10.38	9.91	10.81
20%	0.5	10.37	9.98	10.43
	2	10.38	8.82	10.64
	4	10.40	7.23	10.91

Increasing the *h* parameter has the intended effect of making horizontal inequality larger for a given policy. The value of the bandwidth parameters does not matter as much. As noted earlier, at larger bandwidth our attempt to measure horizontal inequality gets confounded with vertical inequality. That is, with a 20% bandwidth, per-capita refunds tend to equalize incomes in the window within which we are looking for horizontal inequality—and welfare rises.

The bigger issue, as noted earlier, is that it is hard to know what to make of the estimates. In their application, they examine how horizontal equity varies across income and time—so the scale is somewhat irrelevant. For our application, we want to know something about the magnitude of the variation relative to the overall cost. That is, if costs average $10 per household (as they do), is there variation of $1, $10, or $100 around that average cost? We know from the earlier figures that incomes are in fact changing by tens or even hundreds of dollars. Yet, the magnitude of the welfare change is a tiny fraction. As noted earlier, this stems from the fact that this horizontal equity measure is governed by the fractional change in income, which is quite small.

We have a similar issue when we turn to the Slesnick approach. We construct his welfare measures for several values of ρ which is his only parameter in the table below. This yields a welfare effect expressed as a dollar value (note we have not yet implemented the log utility approach). For example, the second row takes the average per capita cost of $10.37 and adds the standard deviation of household costs.

Welfare costs ($ per capita) for alternative policies using Slesnick approach, with $u = y$ (e.g., linear utility). Per capita costs are $10.37 without any inequality effects.

Table 5: Welfare calculations using Slesnick framework

ρ:	*Tradable performance standard*	*Emissions trading (per capita rebate)*	*Emissions trading (income-based rebate)*
0	16.68	73.50	64.29
1	18.80	94.74	83.68
3	23.85	145.21	126.74

With linear utility, it is possible to consider alternative weight on the second, inequality term in the welfare expression. With non-linear utility, the weight has to be sufficient to imply that moving away from the case where all changes in utility are equal definitely reduces welfare and raises costs. As we noted earlier, it might be desirable to allow aversion to inequality to vary continuously towards zero. For that reason, we consider our own measure in the table below,

$$W = \ln\left(1 + \frac{\sum \Delta y_i}{\sum y_i}\right) - \gamma\left(\sum \left|\ln\left(1 + \frac{\Delta y_i}{y_i}\right) - \overline{\ln\left(1 + \frac{\Delta y_i}{y_i}\right)}\right|^{1+\rho}\right)^{\frac{1}{1+\rho}}$$

Now γ can range freely. In fact, we can ask how much weight need we place on inequality to offset a particular difference in $\ln\frac{\sum \Delta y_i}{\sum y_i}$. Below we simply report the value of the two terms for each policy. When $\rho = 0$, we are reporting the average absolute deviation. When $\rho = 1$, we are reporting the standard deviation of the log change. As ρ gets larger, we are tending towards the largest absolute deviation.

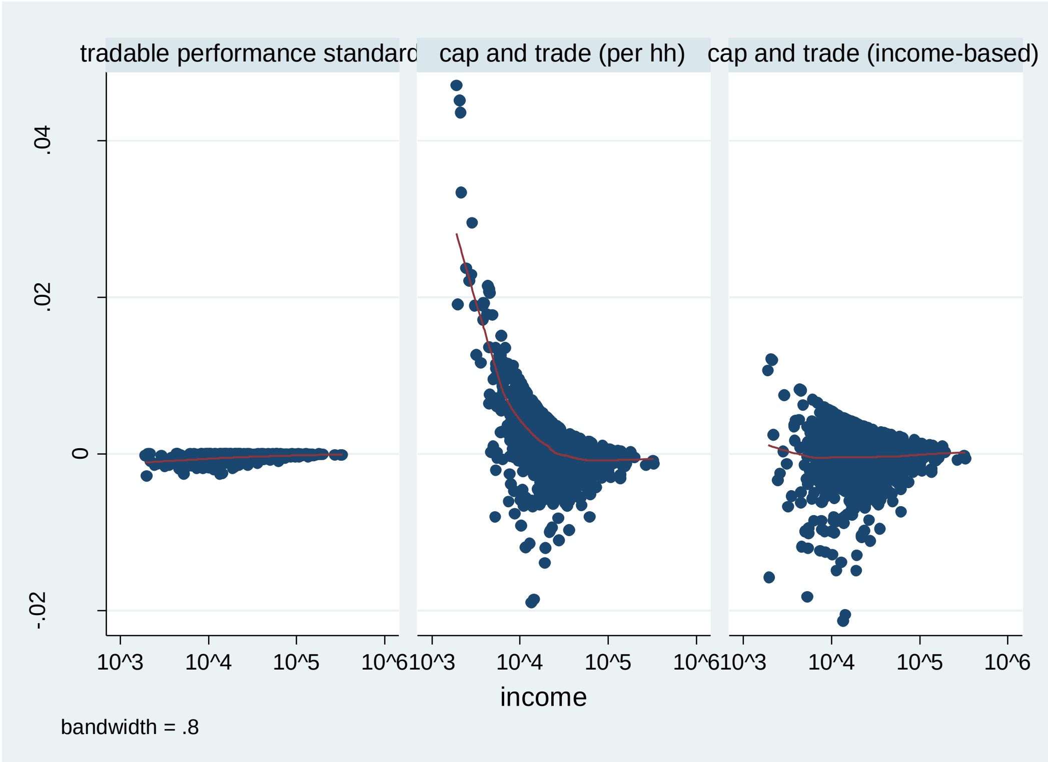
Value of first and second term in Fischer-Pizer welfare measure (all entries can be interpreted as basis points (0.01%) changes in household income.

Table 6: Welfare calculations with Fischer-Pizer framework

	Tradable performance standard	Emissions trading (per capita rebate)	Emissions trading (income-based rebate)
$\ln\left(1+\frac{\sum \Delta y_i}{\sum y_i}\right)$	2.89	2.89	2.89
$\left(\sum \left\|\ln\left(1+\frac{\Delta y_i}{y_i}\right) - \overline{\ln\left(1+\frac{\Delta y_i}{y_i}\right)}\right\|^{1+\rho}\right)^{\frac{1}{1+\rho}}$			
$\rho = 0$	2.49	30.26	21.83
$\rho = 1$	3.58	50.06	32.18
$\rho = 3$	6.52	108.91	57.03

The first line is telling us that the average aggregate costs work out to be about 0.03% of total income. This does not change across policies. The other lines tell us how much the % change in income varies across households. The average absolute deviation (again in basis points) is for $\rho = 0$. The standard deviation, perhaps the most intuitive, is row 2. The bottom row is more heavily weighted towards extreme values. Note that the most extreme values are 4% (400 basis points) in the case of the per capita rebate and 2% (200 basis points for the income-based rebate). Here, at least, there is a more intuitive interpretation. One could compute similar statistics by decile, rather than for the population as a whole.

Similar information is presented graphically below. For each of the three policies, we plot the change in share (essentially the change in logged income) against income on a log scale. We can see the extreme values (4% for the per capita rebate policy and 2% for the income-based rebate). We can also see how the shares and horizontal equity vary across income. In the case of the per capita rebate, some of the variability in shares is coming from a clear redistribution from rich to poor that would not come out in the above statistics alone.



5. Concluding thoughts

We have explored a variety of welfare measures from the literature that could be appropriate for examining energy regulation, particularly as they create heterogeneous effects across households of the same income. Ultimate, the question of equity is not something to be answered by analysis alone, it requires value judgement. However, it is difficult to value something that is not measured, and often attempts to summarize impacts across income groups or other demographic categories ignores important variation within those categories.

For that reason, most of the welfare measures from the literature come up short. It is hard to relate the underlying elements to intuitive concepts that can be presented either without choosing ethical parameters or where the ethical parameters themselves have intuition. The Slesnick measure comes somewhat close, but is constrained. We propose an alternative welfare metric that comes close to replicating the kinds of information that analysts typically present anyway—share effects by income class. What we suggest adding is some measure of the spread, expressible in share terms. Either the standard deviation, average absolute deviation, or something tending towards the maximum absolute deviation. This information, presented either in tables or graphs by income group, could be an important input to both making and understanding policy decisions.

References

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

Fischer, Carolyn, and Alan K. Fox. 2007. "Output-Based Allocation of Emissions Permits for Mitigating Tax and Trade Interactions." *Land Economics* 83 (4): 575–99. doi:10.3368/le.83.4.575.

Jorgenson, Dale W., and Daniel T. Slesnick. 1984. "Aggregate Consumer Behaviour and the Measurement of Inequality." *The Review of Economic Studies* 51 (3): 369–92. doi:10.2307/2297429.

Kaplow, Louis. 1989. "HORIZONTAL EQUITY: MEASURES IN SEARCH OF A PRINCIPLE." *National Tax Journal* 42 (2): 139–54.

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

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books,.

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. doi:10.1016/j.eneco.2011.07.023.

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

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

6. Appendix

We can rewrite the Auerbach and Hassett formula as

$$W = \left[\sum_i N_i \left(\frac{1}{J_i} \sum_j y_{ij}^{1-h} \right)^{\frac{1-v}{1-h}} \right]^{\frac{1}{1-v}} = \left[\sum_i N_i (\tilde{y}_i)^{\,1-v} \right]^{\frac{1}{1-v}}$$

where

$$\tilde{y}_i = \left(\frac{1}{N_i} \sum_j y_{ij}^{1-h} \right)^{\frac{1}{1-h}} = \bar{y}_i \left(\frac{1}{N_i} \sum_j \left(\frac{y_{ij}}{\bar{y}_i} \right)^{1-h} \right)^{\frac{1}{1-h}} = H_i \bar{y}_i$$

and H_i measures how much each income group i is adversely affected by horizontal inequality.

To consider a range of income around each observation, Auerbach and Hassett rewrite their formula as

$$W = \left[\sum_i \left(\sum_j w_{ji} \right) \left(\frac{1}{\sum_j w_{ji}} \sum_j w_{ji} y_j^{1-\gamma} \right)^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}} = \left[\sum_i w_i (\tilde{y}_i)^{\,1-\rho} \right]^{\frac{1}{1-\rho}}$$

where $w_i = \sum_j w_{ji}$ and

$$\tilde{y}_i = y_i \left(\frac{1}{\sum_j w_{ji}} \sum_j w_{ji} \left(y_j / y_i \right)^{1-\gamma} \right)^{\frac{1}{1-\gamma}} = H_i y_i$$

The weights w_{ij} are chosen much like a kernel, to include and weight more heavily observations that are nearby to observation *i* but not too far. They also chosen so that $\sum_i w_{ji} = 1$. That is, the weights are essentially sharing out each observation *i* into adjacent "bins". Each bin, one for each original observation, then becomes a mix of pieces of different observations. The welfare function becomes a weighted average of the adjusted income in each bin, where the weights are just the sum of the fractional weights of each original observation.