The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards

Koichiro Ito
Boston University and NBER

James M. Sallee*
University of Chicago and NBER

January 19, 2014

Abstract

In many countries, fuel-economy standards mandate that vehicles meet a certain fuel economy, but heavier or larger vehicles are allowed to meet a lower standard. This has the perverse implication of allowing automakers to meet standards either by improving fuel economy or by increasing weight, which lowers fuel economy and increases externalities related to accidents. This is but one example of an attribute-based regulation, in which the regulation imposed on a product depends on both the externality it creates and some other attribute. Attribute-basing potentially motivates firms and individuals to strategically alter the attribute, thereby endogenously altering the stringency of the regulation. This paper develops a theory of attribute-basing to demonstrate the costs and benefits of its use. The paper then empirically examines the consequences of attribute-based fuel-economy standards in Japan, where fuel-economy standards are a notched attribute-based function of vehicle weight. We use cross-sectional and panel techniques to demonstrate that attribute-based regulation has significantly altered the distribution of vehicle weight in Japan, where 10% of vehicles bunch at weight notches. For cars whose weight is altered in response to the policy, we estimate that the alteration generates a welfare loss from the exacerbation of weight-related externalities of $1525 per unit sold, which translates into a $686 million annual loss across the Japanese auto market.

Keywords: fuel economy standards, corrective taxation, notches
JEL: H23, Q48, L62

*Ito: ito@bu.edu. Sallee: sallee@uchicago.edu. The authors would like to thank Kunihito Sasaki for excellent research assistance. For helpful comments, we thank Soren Anderson, Damon Jones, Richard Newell, Mar Reguant and seminar participants at the ASSA meetings, Michigan State University, Resources for the Future, and the University of Chicago.

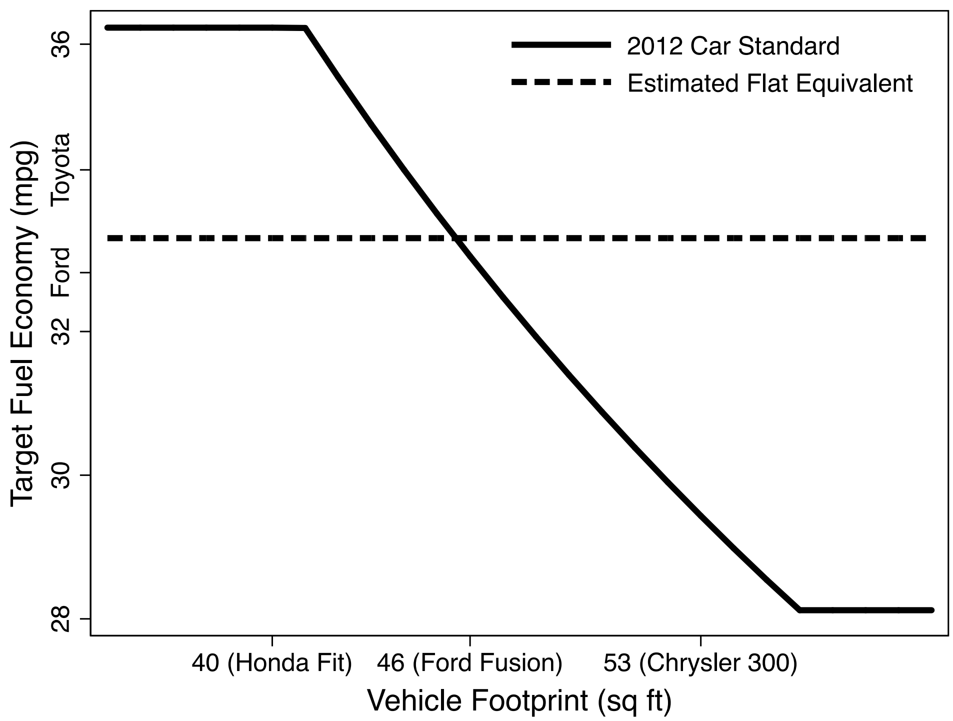
1 Introduction

The goal of this paper is to advance economists' understanding of attribute-basing. Attribute-basing is a feature of many policies aimed at correcting externalities, but it appears to be at odds with the basic logic of Pigouvian taxation. Attribute-based policies are policies in which the corrective policy—which could be a regulation, a tax or a subsidy—imposed on a product is not simply a function of how much externality the product creates, but instead depends on how much externality the product creates relative to some benchmark, which itself depends on some other attribute of the product. Attribute-basing is common to corrective policies, in particular policies aimed at improving the energy efficiency of energy consuming durable goods, but there is no academic literature that explores the optimality, merits and problems with such policies.

Attribute-basing is perhaps best explained by example. Corporate Average Fuel Economy (CAFE) standards in the United States became attribute-based as of 2012. Prior to 2012, every automaker had to sell vehicles that met or exceeded a fleet-average fuel economy of some particular level, which was common to all producers. Since 2012, CAFE specifies a schedule of fuel economy targets for each vehicle that is a declining function of the vehicle's "footprint"—the rectangular area inside of the vehicle's tires. Each automaker has to meet a standard equal to the average target of the vehicles that it sells, so that automakers who sell larger vehicles on average are permitted to have a less efficient fleet. Automakers who exceed their standard are permitted to sell their excess credits to other automakers, so that the policy ultimately imposes a single minimum on the entire market that depends on the entire market's average footprint.

Figure 1 illustrates this schedule for 2012 and shows the large difference in targets across vehicles and automakers. The Honda Fit (a compact car) has a target fuel economy of 36.2 miles per gallon, whereas the Chrysler 300 (a full-size sedan) has a target of 29.4. The steepness of this schedule creates notable differences in the fuel economy required of different automakers who produce cars of different average size. Ford and Toyota, whose targets differ by 1.4 miles-per-gallon in 2012, are labeled in the figure. Vehicle footprint is not immutable, however, so the footprint-based standard not only creates dispersion in requirements across automakers, but also creates an incentive for automakers to enlarge footprint, which runs counter to the goals of the policy by endorsing larger (and hence less energy efficient) vehicles.

Figure 1: Example of an attribute-based standard: 2012 U.S. new car CAFE



As of 2012, the CAFE target fuel economy for a vehicle depends on its footprint. Three footprint values are labeled with example vehicles that have that footprint. The flat equivalent is the EPA's estimate based on the distribution of footprint in current fleet. Prior to 2011, the new car CAFE standard was a flat standard at 27.5 mpg. The firm specific target values for Ford and Toyota based on model year 2011 average footprint is labeled. Data taken from the Federal Register, 40 CFR Part 85.

Such attribute-basing is ubiquitous in the realm of energy efficiency policies. Appliance standards are typically a function of the size or features of the appliance. The same is true for the cutoff value to qualify an appliance for Energy Star certification, which is frequently also used as the criteria for appliance rebate programs meant to conserve energy. The hybrid vehicle tax credit in the U.S. was attribute-based because the subsidy amount was based on each hybrid's gasoline conservation relative to a benchmark vehicle, determined by its market segment. The U.S. is not alone in having attribute-based standards for automobiles. The EU is currently phasing in its first legally binding fuel-economy standards, which has a similar attribute scheme that uses weight rather than footprint. Japan does the same. Chinese fuel economy standards are an attributed-based function of engine displacement. Label ratings, which identify a vehicle on a scale of most to least polluting in various places, including Europe and Korea, depend on vehicle size. All of these policies provide a *lower* standard for products that are larger and *more polluting*, which generates perverse incentives.[1]

A common feature of attribute-based policies is that the slopes of the target function are drawn to fit data. In the U.S., to determine the footprint-target function in the reformed CAFE, government officials used engineering information to estimate the footprint and fuel economy of a fleet of vehicles that used a suite of technological improvements deemed to be feasible in the near future. They then fit a line to the data to determine how quickly fuel economy declined as footprint rose and used this estimated slope as the slope of the attribute-based target. In Japan, the formulation is somewhat different, but it has similar implications. When considering a tightening of fuel economy standards, the Japanese government defines weight categories. Within each weight category it identifies the vehicle with the highest fuel economy. It then determines the required fuel economy rating for vehicles in that weight category in some future target year by multiplying the fuel economy of this "front runner" by some factor, requiring a percentage improvement over the leader in the current period.

In spite of the frequent appearance of attribute-basing in actual policies, the academic literature has provided little scrutiny of its potential costs and benefits. To the extent that these policies are aimed at correcting externalities related to energy consumption, attribute-basing would seem to have no role in an optimal policy because the first-best Pigouvian tax on energy consumption does not depend on anything other than marginal damages. Thus, on its face, attribute-based regulation appears to be superfluous in terms of achieving efficient changes in energy consumption, and it creates an undesirable byproduct via the distortion of the choice of the attribute itself.

Might there be additional considerations that make attribute-basing a good idea? If so, what would characterize the optimal design of an attribute-based policy? If not, so that attribute-basing is inefficient, what distortions does it create and what determines the size of any associated deadweight loss? Is there empirical evidence that these distortions exist? Are they large?

This paper aims to make an important first step in the analysis of attribute-basing by proposing answers to these questions and generally evolving our understanding of this class of public policy.

[1] Counterexamples include circulation and registration taxes for vehicles in Europe. In several countries, these taxes are an increasing function of vehicle characteristics, like cylinders or weight, that are positively correlated with pollution.

We begin in section 2 by establishing an analytical framework that considers two distinct cases, which differ in whether or not the attribute-based regulation allows compliance trading. For policies that allow compliance trading—that is, firms that exceed their standard can sell their excess compliance to other firms who fall short—a regulation imposes a uniform shadow price on all market actors. An example of such a policy is CAFE. In this case, attribute-based regulation is identical to a uniform corrective tax, and the logic of Pigou prevails: attribute-basing is unnecessary and distortionary.

The use of attribute-based regulation in this case creates several distortions. First, it creates a direct distortion analogous to a Harberger triangle in the choice of the attribute. It may also induce an indirect distortion in the choice of energy efficiency. The existence of an attribute-based target causes the second-best tax rate on energy consumption to be below the first-best Pigouvian rate, with the difference between the first and second-best taxes depending on the proportion of compliance that comes from raising the attribute versus improving energy efficiency. Attribute-basing creates additional distortions when the attribute itself is associated with an unpriced externality. This applies to the case of automobiles when the attribute is weight or footprint because those characteristics are correlated with safety externalities (Anderson and Auffhammer Forthcoming).

In the second case, when compliance trading is not allowed, all of the same distortions exist, but there are potential benefits from attribute-basing. If trading is not allowed, then a uniform regulation that forces all products to meet the same energy efficiency minima will generate difference costs of compliance. Attribute-basing can reduce the dispersion in marginal costs to the extent that marginal costs are correlated with the attribute. The optimal attribute slope will generally be less steep than the observed relationship between the attribute and the externality because it represents a compromise between equalizing marginal costs of compliance (which calls for a steeper slope) and distortions in the choice of the attribute induced by the implicit subsidy (which calls for a flatter slope). Real world slopes that were drawn to fit data are the second-best solution only when the attribute is exogenous and is not therefore distorted by the implicit subsidy. Even in this case, however, attribute-basing is less efficient than trading because attribute-basing cannot fully equalize marginal compliance costs unless the attribute perfectly predicts the externality.

The paper then focuses on empirical evidence from Japanese fuel-economy standards, which are particularly amenable for analysis, in section 3. Japanese fuel economy standards require that automakers meet a minimum fuel economy standard, but the standard depends on each vehicle's weight, with heavier vehicles needing to meet laxer standards. Moreover, the minima are discrete functions of weight class, so that vehicles on either side of a weight threshold are required to meet discretely different standards. These "notches" provide a straightforward way to identify distortions in the attribute in response to policy.

Cross-sectional data provide clear evidence of bunching in the weight distribution. We estimate that around 10% of Japanese vehicles increase their weight in response to the policy, with the average weight increase being about 100 kilograms for the effected vehicles. We corroborate these estimates with panel data, and we also estimate an elasticity of the change in weight with respect to the fuel economy standard's stringency and conclude that a one percent increase in the stringency of the standard leads to a .16 percent increase in vehicle weight. Relating these estimates to our theoretical results, we conclude that the second-best subsidy on fuel economy in Japan is 40% lower than the first-best subsidy due to these weight distortions. We also estimate a lower bound on the welfare distortions from the existing policy by quantifying the first-order distortion due to exacerbation of unpriced safety externalities. Based on recent estimates of the value of externalities related to vehicle weight, we estimate that the attribute-basing in the Japanese fuel-economy standards creates welfare losses on the order of $1525 per effected car sold, which aggregates to roughly $686 million annually across the new car market in Japan.

To the best of our knowledge, we are the first paper to take an analytical approach to assessing the merits and distortions caused by attribute-basing in policies aimed at correcting externalities. Our work is nevertheless related to several strands of literature. The literature on the move to footprint-based CAFE rules in the U.S. includes Whitefoot and Skerlos (2012), which uses engineering estimates of design costs and a discrete-choice economic model to predict how much automakers will manipulate footprint in response to a tightening of CAFE standards. That study concludes that an increase in footprint will be a major source of adjustment, which is broadly consistent with our findings of considerable manipulation of weight in the Japanese car market. Gillingham (2013) discusses the implicit incentive for the expansion of footprint in a broader discussion of CAFE policies. Jacbosen (2013) addresses the safety impacts of footprint-based standards in the U.S., which our work will also touch on. Our consideration of notched attributed-based policies relates to the literature on notched corrective taxation, which began with Blinder and Rosen (1985), and includes prior analysis of automobile fuel economy in Sallee and Slemrod (2012).

Our focus in this paper is on attribute-basing in regulations aimed at the energy efficiency of durable goods. Attribute-basing is, however, found in many areas of policy. For example, the liability regime that firms face for worker safety depend on firm size. Similarly, some provisions of the Affordable Care Act only apply to firms of a certain size. Unemployment insurance tax rates depend not only a firm's unemployment history, but also their size and sector. To the extent that these and other policies are intended to correct for an externality, our framework may prove useful in understanding the welfare properties of these types of policies as well.

2 Theoretical framework

The theoretical portion of this paper explores the use of attribute-based regulations for correcting externalities. We first present our notation and define our setup. In section 2.1 we back up and discuss the economic content of our key modeling assumptions before proceeding to the model's solutions.

We model a consumer who buys a single unit of a durable good that has two characteristics, one that is related to an externality (e.g., energy efficiency), denoted e, and some other attribute (e.g., footprint), denoted a. The consumer pays price $P(a, e)$ for one unit of the durable and spends the remainder of his/her exogenous income I on an outside numeraire good x, the price of which we normalize to 1.

Consumer utility is determined by a flow utility from the durable $U(a, e)$, the utility from consumption of all other goods, denoted $h(x)$, and an externality $\phi(E)$, where E is the sum of e over all consumers. We are not particularly interested in complementarities of consumption between the energy-consuming durable and the remainder of the consumer's consumption bundle (or the externality), so we will assume that utility is separable in the three components. Consumer welfare is thus $W = U(a, e) + h(x) + \phi(E)$.[2]

Given our assumption of quasilinearity, the marginal utility of income will be the consumption value of the numeraire h'. To simplify notation, we assume that there is a constant marginal utility

[2]Note that it is reasonable to model energy efficiency as effecting utility only through cost savings that enable greater consumption of the numeraire. We use the more general form here in the hopes that our model be useful beyond the case when e denotes energy efficiency.

of income and we scale utility by this value. Likewise, for notational convenience, we assume that the externality has a constant marginal damage, so that $\phi(E) = \phi \cdot E$. We then rewrite consumer welfare as $W = U(a, e) + x + \phi \cdot E$.

The government can levy a tax or subsidy on the durable that is a function of both attributes, denoted $t(a, e)$. For energy efficiency policies, we have in mind that e is a positive externality, so we write $t(a, e)$ as a subsidy. Our model will focus on the case with a linear tax on e and an attribute-based target function, which implies that $t(\cdot)$ will have the form $t(e - \sigma(a))$, where $\sigma(a)$ is the attribute target function.

For the policies we have in mind, the derivative of $\sigma(a)$, denoted σ', will be negative—as a increases, the target falls and, holding e constant, the subsidy will rise. Figure 1 illustrates the target function for the case of CAFE. In that policy, the target for the Honda Fit is 36.2 mpg. Suppose that the subsidy rate were \$1000 per mpg. If the Fit was rated at 38.2 mpg, it would get a subsidy of $t(e - \sigma(a)) = \$1000(38.2 - 36.2)$. If the Fit was rated at 36.2 mpg, it would get a zero subsidy, and if it were rated at 35.2, it would get a \$1000 tax. Regardless of its starting fuel economy, an increase in fuel economy of 1 mpg would increase the subsidy by \$1000, and an increase of one unit in its footprint would raise the subsidy by an amount equal to the slope of the target function times \$1000.

In our model, all revenue is recycled through a lump-sum demogrant G. In all cases we assume that there are many individual consumers and each therefore ignores their own role in creating the externality—i.e., they take E as fixed. The consumer's optimization problem is then:

$$\begin{aligned} \max_{a,e,x} W =& U(a, e) + x + \phi \cdot E \\ \text{s.t. } & I + G \geq P(a, e) - t(a, e) + x. \end{aligned}$$

Note that this differs from a standard consumer choice problem over the three goods a, e and x only in that the price and tax paid for a and e are possibly related to each other and therefore nonlinear. This is the complete representation of the demand side of the market.

On the supply side, we assume perfect competition. Specifically, we assume that there is a constant marginal cost of the quantity produced of a good with attribute bundle (a, e), which we denote as $C(a, e)$, and that there are no fixed costs or barriers to entry. This means that (in the absence of policy) a consumer will be able to purchase any bundle of attributes (a, e) and pay $P(a, e) = C(a, e)$. Furthermore, because there are no profits in equilibrium due to free entry and the lack of fixed costs, firms will make no profits and consumer surplus will be a sufficient welfare statistic. We assume that $C(a, e)$ is rising and convex in both attributes. In all real world policies that we know of, the attribute a—when written as a good as opposed to a bad—is negatively related to energy efficiency e. (For cars, a is either footprint or weight, both of which correlate negatively with fuel economy.) Thus, we assume that $\partial^2 C/\partial a \partial e \geq 0$. This ensures the convexity of the budget set.[3]

2.1 Discussion of the model's key assumptions

Before discussing the model's solutions, we briefly discuss several key modeling assumptions that we employ, namely the relationship between e and the externality, perfect competition, and the use of a tax policy instead of a regulation.

First, our model assumes that the externality is directly (and linearly) related to the energy efficiency, e. In reality, externalities from energy consumption depend on energy efficiency and utilization, which is itself a function of efficiency. (For example, externalities related to climate change from gasoline consumption depend on vehicle fuel economy and miles driven, and miles driven depends on fuel economy.) This induces what is commonly referred to as a rebound effect.[4] Moreover, policies that target only the new product market induce offsetting effects in the used durable market.[5] For these reasons, it is well understood that energy efficiency policies—whether they are regulatory policies like CAFE or equivalent taxes on the products—represent significant departures from a first-best Pigouvian tax on the externality itself.[6] Our focus is on the marginal efficiency impacts of attribute-basing. For that reason, we will abstract from these considerable concerns and assume that there is some externality related directly to energy efficiency e and that a tax on that feature of a product represents a first-best solution.

Second, our assumption of perfectly competitive supply abstracts from potentially interesting

[3]Empirical estimates in Knittel (2011) indicate that the log of fuel economy is concave down in the log of weight, which is the specific application we have in mind.

[4]The margin of adjustment described here is sometimes called the "direct rebound" effect. There are also other effects that work through income. See Borenstein (2013) for a recent discussion.

[5]See Jacobsen and van Benthem (2013) for an exploration of this effect for the case of automobiles.

[6]See Anderson, Parry, Sallee, and Fischer (2011) for a recent review of CAFE that describes related evidence. For the equivalence of vehicle taxes and CAFE, see Sallee (2011) and Gillingham (2013).

questions regarding how the presence of attribute-based incentives influences welfare through its impact on mark-ups, product differentiation and competition among firms. Despite these limitations, we assume perfect competition to facilitate the comparison between attribute-basing and Pigouvian taxation and to make the welfare analysis transparent and tractable. We believe that our approach will provide a useful foundation upon which to build a subsequent study of attribute-basing that focuses on imperfect competition. Moreover, some studies of CAFE take a short-run perspective, in which the set of vehicles sold is fixed and firms must improve fuel economy by modifying the sales shares of their models. In this sense, perfect competition naturally provides an analysis of the long-run, in which no product characteristics are fixed.

Third, many of our examples regard regulations that, like CAFE, apply to an average across multiple products sold by the same firm. The standard way of modeling firms with multiple products under imperfect competition implies that firms must worry about how increasing some attribute of one of their products will affect the demand for all of their other products. Our assumption about perfect competition abstracts from those concerns because when firms make zero profits and consumers can purchase any bundle of attributes there will be no lost profits from cannibalization. Thus, under perfect competition, nothing is lost by assuming that firms produce a single product, which is the viewpoint we adopt.

Fourth, we model a tax policy instead of a regulation. In appendix A we demonstrate that we can write a tax policy that produces outcomes equivalent to a regulation that requires e to exceed a minimum value of σ. Call the shadow price of the regulation λ. A tax on e that equals $t(e-k)$ is equivalent to the regulation when $t=\lambda$ and $k=\sigma$.[7] Thus, to translate our results, we interpret our tax rate as the shadow price in the equivalent regulation. In appendix A we show that the equivalence extends to attribute-based policies. If the attribute-based regulation requires a product's e to exceed $\sigma(a)$ and this yields a shadow price of λ, then a tax of $t(\omega(a)-k)$ yields the same allocation as the regulation when $\sigma(a)=\omega(a)$ and $t=\lambda$. (These results are true for multi-product or single-product firms.) Given this equivalence, we focus primarily on tax policies, which link more directly to the public finance literature's results on distortionary wedges and Pigouvian

[7] Whenever $k \neq 0$, regulatory policies differ from the Pigouvian tax in that they provide different incentives on the extensive margin. That margin is assumed away in this context—every consumer buys one durable good, though it has been shown to be relevant in other settings, for example in Holland, Hughes, and Knittel (2009). This margin is why taxes on a product's energy consumption are not equivalent to subsidies for its energy efficiency (Metcalf 2009).

taxation.

Fifth and final is our assumption regarding trading. Some regulations, including CAFE, allow firms to trade excess compliance credits across products or firms. If Toyota's fuel economy exceeds its target, for example, it may sell its excess compliance to BMW, who can credit their account accordingly. When full trading exists, the shadow prices of the regulation will be equalized across all producers. In that case, the equivalent tax model imposes a single tax rate on all firms. In some other policies, there is no trading. Appliance standards, for example, generally require each product to meet the regulatory minimum individually. In this case, the individual products may face different shadow prices, and the equivalent tax policy would be one in which tax rates differ across products. We will consider each of these cases separately, starting with the case of trading.

2.2 Baseline model with trading

We begin by considering the case with trading, so that there is a single shadow price (tax rate) facing all consumers. In this case, we focus on the problem with only one type of consumer, which provides great notational simplification. (If there were many types of consumers, their problems would be separable because of our perfect competition assumptions and our solutions would differ only in featuring summations over types.) We further assume that there are measure 1 consumers, so that $E = e$.

2.2.1 First-best solutions do not involve attribute-basing

The first step in our analysis is to derive the first-order conditions for the consumer facing a tax policy and compare them to the planner's first-best conditions when the planner is allowed to choose the allocation directly. We can then see what function $t(a, e)$ would make the consumer's conditions equal to the first-best.

Under these assumptions, we now write the consumer's problem as an unconstrained optimization problem by substituting the budget constraint into the objective function:[8]

$$\max_{a,e} W = U(a, e) + [I + G - P(a, e) + t \cdot (e - \sigma(a))] + \phi \cdot E.$$

[8]Our assumptions regarding $C(a, e)$ ensure that the budget set is convex and thus that the consumer will exhaust their budget.

This is a standard choice problem, and the first-order conditions can be derived directly by differentiation:

$$\frac{\partial W}{\partial a} = \frac{\partial U}{\partial a} - \frac{\partial P}{\partial a} - t\sigma' = 0 \tag{1}$$
$$\frac{\partial W}{\partial e} = \frac{\partial U}{\partial e} - \frac{\partial P}{\partial e} + t = 0, \tag{2}$$

where σ' is the derivative of $\sigma(a)$ with respect to a.

First-order conditions for the first-best allocation can be found by differentiating the planner's problem, which is:

$$\max_{a,e} SWF = U(a,e) + [I - C(a,e)] + \phi \cdot e.$$

This differs from the consumer's problem in directly inserting the cost function in place of prices (though these will be equal under perfect competition) and taxes, and in recognizing how the choice of e affects the total externality. The first-best optimization conditions are found by differentiation:

$$\frac{\partial SWF}{\partial a} = \frac{\partial U}{\partial a} - \frac{\partial C}{\partial a} = 0 \tag{3}$$
$$\frac{\partial SWF}{\partial e} = \frac{\partial U}{\partial e} - \frac{\partial C}{\partial e} + \phi = 0. \tag{4}$$

Under perfect competition, prices reflect unit costs, so $\frac{\partial P}{\partial a} = \frac{\partial C}{\partial a}$ and $\frac{\partial P}{\partial e} = \frac{\partial C}{\partial e}$. Then, it is immediately apparent that the only way to make the consumer's first-order conditions equal to the first-best is to set $\sigma' = 0$ and $t = \phi$. That is, to have a Pigouvian subsidy on e with no attribute-basing.

This is the first result of our theoretical analysis. The result is not surprising, but it is relevant for evaluating policy, as it establishes a set of conditions under which attribute-basing is counterproductive. When there is compliance trading and the externality associated with e is the only market failure, then the optimal policy will not involve attribute-basing. The optimal attribute-based slope is uniquely zero. This result will be unchanged if there are multiple types of consumers. If there is a role for attribute-basing, it must be due to some additional factor. We discuss several possibilities below.

2.2.2 Mispricing of externality does not justify attribute basing

The analysis above focuses on the first-best allocation, in which attribute-basing has no role to play. One possible explanation for real world attribute-based policies is that policymakers are unable to set the stringency of the regulation (or the size of the tax or subsidy) to the first-best level, due to political constraints. Might this create a role for attribute-basing? In terms of the model, what we have in mind is that the policymaker is forced to set a too low subsidy, so that $t < \phi$.

To see whether this justifies attribute-basing, we can derive the second-best σ', taking $t < \phi$ as given. We focus on the linear case here, where $\sigma(a) = \sigma' \cdot a + \kappa$, so that $t(a, e) = t \cdot (e - \sigma' a - \kappa)$. In that case, the second-best attribute slope is characterized by solving the following problem:

$$\max_{\sigma'} U(a, e) - P(a, e) + \phi e.$$

The first-order condition of this problem is:

$$\left(\frac{\partial U}{\partial e} - \frac{\partial P}{\partial e} + \phi\right)\frac{\partial e}{\partial \sigma'} + \left(\frac{\partial U}{\partial a} - \frac{\partial P}{\partial a}\right)\frac{\partial a}{\partial \sigma'} = 0.$$

Into this we substitute equations 1 and 2, the optimality conditions from the consumer's problem, and rearrange. This yields

$$\sigma'^{*} = \underbrace{\frac{t - \phi}{t}}_{-} \frac{\partial e / \partial \sigma'}{\partial a / \partial \sigma'}. \tag{5}$$

To interpret this result, note that the first factor is negative (because $\phi > t > 0$) and $\partial a/\partial \sigma'$ is negative.[9] The sign of $\partial e/\partial \sigma'$ is ambiguous. Changing σ' while holding t constant has no direct effect on the price of e; instead it changes the price, and thus presumably the quantity, of a, which can lead to a change in the choice of e. The direction of this change is ambiguous, and may in fact be zero. As we discuss further below, it is natural to suppose that $\partial e/\partial \sigma' < 0$ because an increase in a raises the marginal price of e, but it is not obvious that the marginal utility of e will change at all. In this case, σ'^{*} will be *positive*, which is the opposite of policy. It is *possible* that $\partial e/\partial \sigma' = 0$, in which case the second-best attribute slope will be zero, or that $\partial e/\partial \sigma'$ is negative, in which case the slope will be negative. Even in this case, however, the second-best slope will bear

[9] There are no income effects, and the implicit subsidy of a is equal to $-t\sigma'$. Thus, $\partial a/\partial \sigma'$ is negative because an increase in σ' will lower the subsidy and thus induce less endogenous adjustment of a.

no relationship to real world slopes that are chosen by fitting data, as described in the introduction. In short, a tax rate that is too small is unlikely to rationalize the attribute-basing that we observe in the world.

There are many real world policies in which it appears that policymakers are constrained by political factors from setting regulations or taxes at the right level. It is tempting to think that the introduction of attribute-basing into regulations would give the policymakers greater flexibility and therefore help circumvent constraints on setting the right level. This is unlikely to be the case, however, because attribute-basing is a very unwieldy tool for affecting the choice of the product characteristic that creates the externality; it primarily acts as an implicit subsidy for the attribute itself, not the externality generating characteristic.

2.2.3 Attribute basing reduces the second-best tax on the externality

An alternative second-best exercise is to assume that the degree of attribute-basing is taken as given—perhaps because it is the result of political bargaining—and ask how this influences the optimal tax rate t. In the linear case, when σ' is fixed, planner's problem is:

$$\max_t U(a, e) - P(a, e) + \phi e.$$

The planner's first-order condition is:

$$\left(\frac{\partial U}{\partial e} - \frac{\partial P}{\partial e} + \phi\right)\frac{\partial e}{\partial t} + \left(\frac{\partial U}{\partial a} - \frac{\partial P}{\partial a}\right)\frac{\partial a}{\partial t} = 0.$$

Substituting in the first-order conditions of the consumer (equations 1 and 2) and rearranging yields:

$$t^* = \frac{\phi}{1 - \sigma'\left(\frac{\partial a}{\partial t} / \frac{\partial e}{\partial t}\right)}. \tag{6}$$

Increasing t directly subsidizes both e and a (assuming that $\sigma' < 0$). Thus, we expect both $\partial a/\partial t$ and $\partial e/\partial t$ to be positive. If so, then the denominator will be greater than one, and the second-best subsidy rate t^* will be *below* ϕ, which is the first-best Pigouvian subsidy that equals marginal damages. The second-best subsidy will equal the Pigouvian benchmark if $\sigma' = 0$ or if a is perfectly inelastic, so that $\partial a/\partial \sigma' = 0$. Otherwise, the second-best subsidy will be below marginal damage, and it will be further below this benchmark as the attribute slope is steeper (the larger is $|\sigma'|$) and as the responsiveness of a, relative to the responsiveness of e, is greater.

This latter point can perhaps be better seen when rewriting equation 6 in terms of elasticities:

$$t^* = \frac{\phi}{1 - \epsilon_a^{\sigma} \frac{\epsilon_t^a}{\epsilon_t^e}}, \tag{7}$$

where ϵ_y^x denotes the elasticity of x with respect to y for each pair of variables.[10] Equation 7 shows that the second-best tax is driven further from the first-best rate when the elasticity of a with respect to the tax rate is large compared to the elasticity of e. That is, the market can respond to the policy by changing both a and e. When the market response is tilted more heavily towards a, this will drive the second-best tax rate towards zero because the response in a is distortionary, whereas the second-best policy will be close to the Pigouvian benchmark when a is not very responsive. This highlights the importance of determining the responsiveness of a to policy, which the empirical portion of our paper does for the case of Japanese fuel-economy standards. To preview our results, we estimate the response of both weight (a) and fuel economy (e) to the tightening of the Japanese standard (t) and find that the response in weight is significant and that, as a consequence, the second-best tax rate is approximately 40% below the Pigouvian rate.

2.2.4 Attribute-basing causes deadweight loss

A final exercise in assessing the consequences of attribute-basing in the presence of compliance trading is to assume that policymakers set the Pigouvian tax $t = \phi$, but that some attribute-basing is used nevertheless, so that $\sigma' \neq 0$. We explore the welfare consequences of this scenario, which might prevail if policymakers introduce attribute-basing for political reasons or out of confusion, but do not understand how its presence changes the optimal choice of t.

A qualitative characterization of distortions can be made by examining the first-order conditions of the consumer, again comparing this to the first-best. Let a^* and e^* denote the first-best choices and a' and e' denote the values chosen under the attribute standard. Using this notation, the

[10] This transformation follows from substitution based on the definition of the elasticities and algebraic simplification, which uses the fact that $e = \sigma(a)$ at the constrained optimum.

planner's and consumer's first-order conditions (equations 2 and 4) with respect to e are:

$$\begin{aligned} \text{First-best:} \qquad & \frac{\partial U(a^*, e^*)}{\partial e} + \phi = \frac{\partial P(a^*, e^*)}{\partial e} \\ \text{Consumer's choice:} \qquad & \frac{\partial U(a', e')}{\partial e} + t = \frac{\partial P(a', e')}{\partial e}. \end{aligned}$$

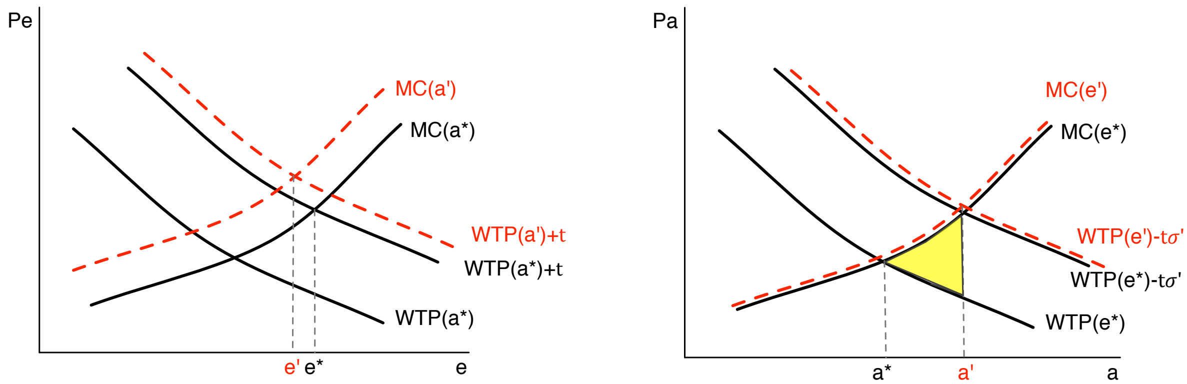
When $t = \phi$ and $\sigma' \neq 0$, the planner's and consumer's first-order conditions are the same, except that the value of a will be different. This means that there are no "partial equilibrium" distortions in the choice of e. Instead, any distortion in the choice of e comes from the fact that the attribute slope provides an inefficient subsidy of a that distorts the choice of a, which may have "general equilibrium" effects on the choice of e.

This is illustrated in Figure 2a, which shows the consumer's choice of e by plotting the marginal cost and marginal benefit of (willingness to pay for) e, conditional on $a = a^*$. (We do not need to condition on x because of the quasi-linearity of $U(a, e)$.) The marginal cost of increasing e to the consumer is the derivative of the cost function with respect to e, given a level of a. The private marginal benefit of increasing e is the derivative of U with respect to e, for a given level of a. The solid black lines plot hypothetical curves for these functions when $a = a^*$. The corrective tax rate t shifts the willingness to pay curve up by amount t. Thus, the consumer will choose e^* when there is no attribute-basing and the tax rate is t.

Suppose that, under attribute-basing, the consumer's choice of a, call it a', exceeds a^*. In the diagram, this will shift both the marginal cost and the marginal benefit curves, depending on the cross-partial derivatives of price and flow utility. The marginal cost curve will shift up by $\approx (a' - a^*)\frac{\partial^2 P(a^*,e^*)}{\partial a \partial e}$. The willingness to pay curve will shift by $\approx (a' - a^*)\frac{\partial^2 U(a^*,e^*)}{\partial a \partial e}$. If $\frac{\partial^2 P(a^*,e^*)}{\partial a \partial e} = \frac{\partial^2 U(a^*,e^*)}{\partial a \partial e}$ (which would occur only by coincidence) then $e^* = e'$ and attribute-basing would cause no change in the allocation of e. Physics suggests that the cross-partial is positive for the cost function, but the sign of the cross-partial for utility is ambiguous. If it is positive, then the two effects will be offsetting. If it is negative or zero, then the general equilibrium effect on e will unambiguously lead to a decrease in e.

Figure 2: Distortions in the choice of a and e from attribute-basing

(a) Choice of e **(b)** Choice of a



The same is not true in the analysis of a. The first-order conditions with respect to a are:

$$\text{First-best:} \quad \frac{\partial U(a^*, e^*)}{\partial a} = \frac{\partial P(a^*, e^*)}{\partial a}$$

$$\text{Consumer's choice:} \quad \frac{\partial U(a', e')}{\partial a} + \underbrace{t\sigma'}_{>0} = \frac{\partial P(a', e')}{\partial a}.$$

When $\sigma' \neq 0$, attribute-basing introduces a wedge between the marginal benefit and marginal cost functions for a. This is directly analogous to a Harberger triangle in standard partial equilibrium tax analysis, where the tax wedge is of size $\sigma' t$. This is shown in Figure 2b, which plots the marginal cost and benefit of a, conditional on a given level of e. The introduction of the attribute-based wedge shifts the benefit curve out linearly with σ'. The deadweight loss due to the price wedge will be $\approx 1/2 \cdot (a' - a^*)\sigma' t$. This approximated loss is "second-order" in the same sense that the canonical Harberger triangle is; it is a function of $(\sigma' t)^2$ because $a' - a^*$ is locally linear in $\sigma' t$.

The difference between e^* and e' will also cause a change in the choice of a. This will be a "general equilibrium" effect, of magnitude $\approx (e' - e^*)\left(\frac{\partial^2 P(a^*,e^*)}{\partial a \partial e} - \frac{\partial^2 U(a^*,e^*)}{\partial a \partial e}\right)$. The direct effect of attribute-based standards is in creating the wedge $t\sigma'$, which equals $\phi\sigma'$ when t is set at the Pigouvian level. This wedge will cause the consumer to choose too high of a value of a, and the distortion will be rising as σ gets steeper (σ' gets more negative). The wedge is also larger (in absolute value) as t gets larger (or as ϕ gets larger assuming the Pigouvian rate for t is used).

This has important implications for empirical studies of attribute-based policies because it implies that the main welfare losses may be due to distortions in the attribute choice, not in the amount of externality reduction that is achieved. Take CAFE for example. The introduction of attribute-basing may have a limited impact on fuel economy as compared to a policy with the same implicit subsidy for fuel economy that has no attribute-basing. Fuel economy might be higher or lower under the attribute-based policy than under the shadow-price equivalent attribute-free policy. The largest distortions and welfare losses of the footprint-based standards may be associated with the distortion in the footprint, not the distortion in fuel economy. The size of the externality enters the distortion formulas because it determines the size of the tax wedge that distorts the attribute, not because the primary welfare loss is from too little externality reduction. Altogether, this suggests that the place to look for distortions in the data may be in the distribution of the attribute upon which standards are based, not in the distribution of energy efficiency. This is what we do below for the case of Japanese fuel-economy standards, where we focus on distortions in the distribution of vehicle weight.

2.2.5 Distortions in a are first-order if a has an unpriced externality

In some real world attribute-based policies, the attribute upon which the target is based itself generates externalities. Take the example of Japanese fuel economy standards, in which case e is fuel economy and a is vehicle weight. Heavier vehicles pose greater risk to drivers of other vehicles in the case of a collision (Anderson and Auffhammer Forthcoming). This risk is unpriced (whereas the increased safety of those in the heavier car itself is a private benefit and therefore should be priced in the market) and represents a negative externality associated with weight.[11]

The model can accommodate this possibility by adding a second externality to the welfare function equal to $\omega \cdot a$, so $W = U(a, e) + x + \phi e - \omega a$, with $\omega > 0$. To see the implications of this for the distortion related to the use of attribute-based standards in the baseline model, we can consider an approximation of the welfare loss due to a movement of choices from the Pigouvian benchmark (the choice when $t = \phi$ and $\sigma' = 0$), denoted (a^*, e^*), to that which is chosen by the consumer under the attribute-based policy, denoted (a', e'). The first-order Taylor expansion for

[11] In principle, the externality risk may be partly priced through insurance or legal liability. White (2004), however, argues that neither tort liability nor mandatory liability insurance prices safety externalities. In brief, tort liability requires negligence, not just that one be driving a dangerous vehicle. Liability insurance generally coves the cost of damages to a vehicle, but it is but a small fraction of the value of a life. In addition, rate differences across vehicles are very coarse and reflect average driver characteristics in concert with vehicle attributes.

this welfare loss, when a has an externality of ω, is:

$$\begin{aligned} W(a', e') - W(a^*, e^*) =& (a' - a^*) \left(\frac{\partial U(a^*, e^*)}{\partial a} - \frac{\partial P(a^*, e^*)}{\partial a} - \omega \right) \\ &+ (e' - e^*) \left(\frac{\partial U(a^*, e^*)}{\partial e} - \frac{\partial P(a^*, e^*)}{\partial e} + \phi \right) \\ =& - (a' - a^*)\omega. \end{aligned} \tag{8}$$

Note that $\frac{\partial U(a^*,e^*)}{\partial a} = \frac{\partial P(a^*,e^*)}{\partial a}$ and $\frac{\partial U(a^*,e^*)}{\partial e} = \frac{\partial P(a^*,e^*)}{\partial e} - \phi$, and several components therefore cancel. (This is just an application of the envelope theorem, which applies at the first-best choice.) What remains is that, to a first-order approximation, the welfare loss from attribute-basing will be equal to the change in a times the externality associated with a. This is an externality "rectangle", which will have area $-(a' - a^*)\omega \approx s \cdot \sigma' t$, where s is the joint supply and demand elasticity of a with respect to the price wedge $\sigma' t$. That is, this externality rectangle is linear in $t\sigma'$, whereas the distortion from private utility lost (the Harberger triangle) is quadratic in $t\sigma'$.

Intuitively, this indicates that when the attribute a has an unpriced externality associated with it, the welfare loss from the introduction of a "small" attribute slope will be dominated by the externality associated with a. This is relevant for our empirical evaluation, where we use this reasoning to justify a focus on how Japanese fuel economy standards cause an increase in average vehicle weight. (Of course, when actual policies are far from the optimum the other terms may dominate.)

If the policymaker were aware of the externality associated with a, the attribute-based policy could be designed to correct both externalities by setting $t = \phi$ and $\sigma' = \omega/\phi > 0$. That is, if there are two externalities, then the attribute-based standard can affect a Pigouvian tax on each. This would imply an attribute slope which has the opposite sign as what we observe in reality, suggesting that policymakers are not using the degree of freedom in the policy to correct this second externality when it exists.

2.2.6 Results are similar when $t(\cdot)$ has notches

In appendix B we explore the welfare implications of notched attribute-functions, that is, cases in which the tax function $t(\cdot)$ may be discrete. The Japanese fuel economy regulation, for example, has a target function that is stepwise in weight (a). The key results of that model is twofold. First, all of the intuition developed above about optimal policy design and the nature of distortions carry over directly to the case with notches. Second, in the case where the tax function is continuous in e but notched in a (like the Japanese fuel economy regulation), all of the distortions come from products that choose to bunch at notch points in a. Products that do not choose a to be exactly at a notch are not distorted because the notched subsidies create only income effects, which do not affect the choice of a or e when utility is quasilinear. This is useful for interpreting our empirical results, and we return to these theoretical results when discussing our empirical results.

2.3 The model when there is no trading

Up to now we have focused on the case where a common tax rate applies to all consumers and firms, which corresponds to a regulatory policy that allows trading. This characterizes a number of regulations, including CAFE and feebates in the automobile market. Some attribute-based regulations, however, apply to each and every product in a market, or they apply to each firm separately and trading is disallowed. Key examples are appliance product standards, which specify a minimum energy efficiency that all products are required to meet. Refrigerators in the U.S., for example, must meet a minimum energy efficiency that is a continuous function of its "adjusted volume" (equal to fresh volume plus 1.63 times freezer volume) and a discontinuous function of features, including door type (french or not), location of freezer (top, bottom or side) and whether there is through the door ice.

When trading is not allowed, products will face different shadow prices if each product is required to meet the same minimum. This is equivalent to a tax policy that imposes different tax rates on e across different products. This will cause the marginal costs of compliance to be different across goods, thereby implying inefficient provision of e. In this setting, attribute-based regulation may provide a benefit by (partially) equalizing shadow prices. This benefit can be weighed against the costs stemming from distortions in the choice of a.

2.4 Second-best policy in the absence of compliance trading

To capture the nature of non-trading, it is more natural to work with an attribute based standard than a tax. And, in this case, it is necessary to consider heterogeneity. We model heterogeneity as a discrete set of types of consumers $n = 1, ..., N$, each of whom chooses a separate vehicle because they have a different flow utility function $U_n(\cdot)$. For notational simplicity, we assume that there are measure 1 total consumers, and each type has measure $1/N$. Given quasilinearity, it is inconsequential to assume that I is common across types, so we make that assumption to streamline notation.

Under a standard, the problem of the type n consumer is:

$$\max_{a_n, e_n} W = U_n(a_n, e_n) + [I + G - P(a_n, e_n)] + \lambda_n \cdot (e_n - \sigma(a_n)) + \phi \cdot E,$$

where λ_n is the shadow price on the regulatory requirement that $e > \sigma(a)$. The consumer's first-order conditions are:

$$\frac{\partial W}{\partial a} = \frac{\partial U}{\partial a} - \frac{\partial P}{\partial a} - \lambda_n \sigma' = 0 \tag{9}$$

$$\frac{\partial W}{\partial e} = \frac{\partial U}{\partial e} - \frac{\partial P}{\partial e} + \lambda_n = 0, \tag{10}$$

Note that these conditions are identical to those in the tax case above (equations 1 and 2) when $\lambda_n = t$.

We assume that the planner solves a utilitarian social welfare function that puts equal weight on utility across types, maximizing with respect to the policy function $\sigma(a)$. We focus here on the case of a linear attribute target function, which we denote as $\sigma(a) = \sigma' a + \kappa$, and on the case in which each product must individually meet the standard (as in the case of appliances). We also assume that the standard is binding for all products, so that $e_n = \sigma' a_n + \kappa$ exactly. The planner's problem can be written as:

$$\max_{\sigma', \kappa} SWF = \sum_n \left(U_n(a_n, e_n) + [I + G - P(a_n, e_n)] + \phi \cdot e_n \right).$$

Note that in the case of a standard, the planner does not choose directly a shadow price (tax rate), but rather chooses one implicitly by setting the stringency of the standard. It is useful to think of κ as determining the overall level of the standard, while σ' determines how large is the concession to products with a larger a.

The first-order condition with respect to κ is:

$$\sum_n \left(\frac{\partial U_n}{\partial a_n} - \frac{\partial P}{\partial a_n}\right)\frac{\partial a_n}{\partial \kappa} + \left(\frac{\partial U_n}{\partial e_n} - \frac{\partial P}{\partial e_n} + \phi\right)\frac{\partial e_n}{\partial \kappa} = 0.$$

Using the optimality conditions from the consumer's problem (equations 9 and 10), this can be rewritten as:

$$\sum_n \lambda_n \sigma' \frac{\partial a_n}{\partial \kappa} + (\phi - \lambda_n)\frac{\partial e_n}{\partial \kappa} = 0. \tag{11}$$

When the constraint is binding, $e_n = \sigma' a_n + \kappa$. Total differentiation of this constraint yields a relationship between $\partial a_n/\partial\kappa$ and $\partial e_n/\partial\kappa$, namely that $\partial e_n/\partial\kappa = \sigma'\partial a_n/\partial\kappa + 1$. Using this substitution and rearranging equation 11 yields:

$$\frac{\sum_n \lambda_n}{n} = \phi\left(1 + \underbrace{\sigma' \frac{\sum_n \frac{\partial a_n}{\partial \kappa}}{n}}_{+}\right), \tag{12}$$

which says that, at the planner's optimum, the *average* shadow price ($n^{-1}\sum_n \lambda_n$) is equal to the marginal externality ϕ times a factor. The factor is one plus the attribute slope times the average responsiveness of a to a tightening of the standard. We expect that $\partial a_n/\partial\kappa$ is positive, so that, with $\sigma' < 0$, the presence of the attribute slope makes the average optimal shadow price *less* than the Pigouvian benchmark that would have $\lambda = \phi$. This result is closely related to the second-best subsidy in the trading case derived in equation 6. The steeper the attribute slope and the more responsive is a (on average) to a tightening of the standard, the weaker will be the second-best standard, relative to the Pigouvian benchmark. In this case, however, it is not optimal to have no attribute sloping.

To see that, consider the first-order condition for σ', which is:

$$\sum_n \left(\frac{\partial U_n}{\partial a_n} - \frac{\partial P}{\partial a_n}\right)\frac{\partial a_n}{\partial \sigma'} + \left(\frac{\partial U_n}{\partial e_n} - \frac{\partial P}{\partial e_n} + \phi\right)\frac{\partial e_n}{\partial \sigma'} = 0.$$

Total differentiation of the constraint yields $\partial e/\partial\sigma' = \sigma'\partial a/\partial\sigma' + a$. Using this substitution and rearranging yields:

$$\frac{\sum_n \lambda_n a_n}{n} = \phi\left(\bar{a_n} + \sigma' \frac{\sum_n \frac{\partial a_n}{\partial \sigma'}}{n}\right), \tag{13}$$

where $\bar{a_n}$ denotes the mean of a_n. We then use the definition of the sample covariance of λ_n and a_n to rewrite $\sum_n \lambda_n a_n$ as $\sum_n(\lambda_n - \bar{\lambda})(a_n - \bar{a}) + n^{-1}\sum_n a_n \sum \lambda_n$, substitute equation 12 and rearrange to get the following result regarding this covariance:

$$\frac{\sum_n(\lambda_n - \bar{\lambda})(a_n - \bar{a})}{n} = \phi\sigma' \left(\underbrace{\frac{\sum_n \frac{\partial a_n}{\partial \sigma'}}{n}}_{-} - \bar{a_n}\underbrace{\frac{\sum_n \frac{\partial a_n}{\partial \kappa}}{n}}_{+}\right). \tag{14}$$

This result expresses the covariance of the marginal shadow price λ_n with the attribute a_n. If a is not responsive to policy, so that $\partial a_n/\partial\sigma' = \partial a_n/\partial\kappa = 0$, then the optimal policy will produce a zero covariance between shadow prices and the attribute. As we discuss further below, this has an intuitive connection to the zero correlation between a regressor and residuals in least squares regression.

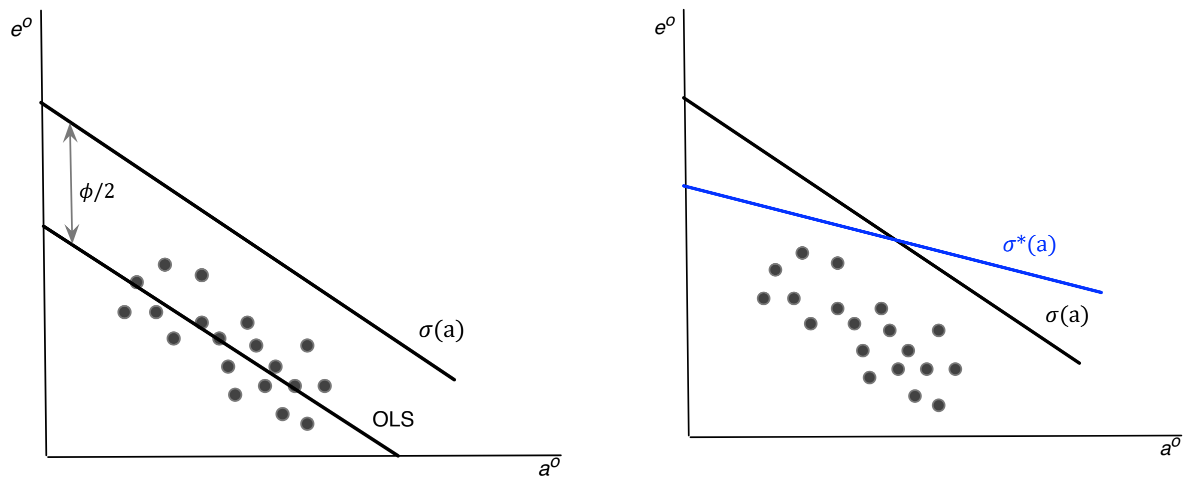
As long as a is endogenous, however, optimal policy will induce a correlation between shadow prices and the attribute, which can be signed. Typically $\partial a_n/\partial\sigma'$ will be negative (a more positive attribute slope leads to less a) and $\partial a/\partial\kappa$ will be positive (a tighter standard leads to more a). Thus, when $\sigma' < 0$, as in our reference policies, the optimal covariance will be positive. This means that products with a larger a will have larger shadow prices, on average. The intuition for this is that the steepness of the attribute slope is chosen to minimize dispersion in shadow prices, which is achieved when shadow prices are uncorrelated with the attribute. But, a steeper attribute slope causes greater distortion in the choice of a. Trading off these distortions is what causes the optimum to feature a correlation between λ and a. This correlation is greater as the responsiveness of a to policy is greater, which parallels our result regarding the second-best policies in the trading case.

To see this intuition more clearly, consider the special case where a is exogenous and when the shadow price of regulation equals $\lambda_n = 2(\sigma' a_n + \kappa - e_n^o)$, where e_n^o is the value of e that consumer n would choose in the absence of policy (it is their private optimum). This special case supposes that all n have a common shadow price function, which depends only on the distance between the

Figure 3: The second-best attribute slope

(a) Special case when a fixed

(b) When a endogenous



privately optimal e_n^o that the consumer would choose in the absence of policy and the value of e that they are forced to choose, which is $\sigma' a_n + \kappa$. In this case, the planner's first-order conditions are:

$$\sum_n 2(e_n^o - \sigma' a_n - \kappa) + \phi = 0 \tag{15}$$

$$\sum_n 2(e_n^o - \sigma' a_n - \kappa) a_n + \phi a_n = 0 \tag{16}$$

If the ϕ terms were missing, these first-order conditions would be identical to the first-order conditions of bivariate ordinary least squares (OLS). The OLS solutions would be $\kappa = \bar{e} - \sigma'\bar{a}$ and $\sigma' = cov(e_n^o, a_n)/var(a_n)$. The externality ϕ means that one does not actually minimize the sum of squared residuals, but instead tries to average them at ϕ. The result is that the optimal slope is unchanged but the optimal intercept is shifted up to $\kappa = \phi/2 + \bar{e} - \sigma'\bar{a}$.[12]

Figure 3a illustrates this intuition. Consider a set of data points for the privately optimal values of a and e, labeled a^o and e^o. OLS would fit a line to the data, as illustrated. If a^o were fixed, then the optimal linear attribute-based target function, denoted simply as $\sigma(a)$ in the figure, would be a line parallel to the OLS fitted line but shifted up by amount $\phi/2$. This is the special case for which

[12]This results in an average value of the change in e from e^o of $\phi/2$, which is the first-order condition when the total shadow price is $(\sigma' a_n + \kappa - e_n^o)^2$.

actual polices are the solution. Fitting a line to the data is only the second-best solution when a is exogenous and not distorted by the policy, which again highlights the importance of determining whether or not a responds to policy, which is the objective of our empirical analysis.

When a is endogenous, however, the optimal target function will induce a positive correlation between the shadow price and the attribute, according to equation 14. This is illustrated in figure 3b, which denotes the second-best attribute function as $\sigma'^{*}(a)$ when a is endogenous, alongside the optimum when a is fixed, denoted $\sigma(a)$. The endogeneity of a lowers the intercept and flattens the slope, as compared to the case with exogenous a. That is, the optimal attribute slope is a compromise between the OLS slope, which seeks to minimize the adjustment costs across all products, and the fact that steeper slopes lead to greater distortions in a. When a responds very little, the optimal slope will be close to that of OLS. The empirical portion of this paper shows that a is highly responsive for the case of Japanese fuel-economy standards, suggesting that, at least for that case, the optimal slope will deviate substantially from the OLS fitted line.

2.4.1 Attribute-basing is inefficient compared to trading

The special case when a is fixed and the shadow cost of compliance is a common function across n that depends only on the difference between the standard and the privately optimal choice of e is the most favorable possible situation for real world attribute-based regulation. Even in this case, however, attribute-basing is inefficient compared to compliance trading because the attribute-based regulation will not completely equalize shadow prices. For the OLS case, when the shadow price $\lambda_n = 2(\sigma' a_n + \kappa - e_n^o)$, the fraction of shadow price equalization that is achieved is the R^2 of the regression of e on a in the product data. If the R^2 of the regression of e_n on a_n is 1, then the linear attribute schedule fits the optimal schedule perfectly, and the policy matches the first-best. But, as the R^2 gets lower, it implies that the average deviation between the actual improvement required of each type, Δe_n, and the optimal amount of $\phi/2$ is getting larger, and therefore inefficiency is greater. In the case of Japan, this R^2 is around .75, suggesting that the attribute slope there captures about three-quarters of the variation in marginal costs (in the special case where all vehicles have equal marginal cost).

The analogy to OLS is, of course, specific to the case where the marginal cost function m is quadratic. With some other loss function, the second-best schedule will have a different fit and a goodness-of-fit statistic other than the R^2 will be the appropriate welfare metric. But the basic intuition will carry over to other loss functions. The key intuition, which is general, is that the efficiency of the attribute-based schedule is tied to the correlation between the optimal value σ_n and the attribute amount a_n. To the extent that a_n predicts σ_n exactly, a fully flexible attribute-based policy will be able to capture a larger share of the welfare gains. In contrast, compliance trading would achieve full efficiency.

The inefficiency of attribute-basing as compared to compliance trading is, however, likely far greater than is implied by the R^2, which applies only to the special case, for two reasons. First, the optimal improvement for each n is likely to differ, because the shadow price function will depend on the curvature of $U_n(\cdot)$, which is likely to differ across consumers or products. That is, to the extent that marginal compliance costs vary conditional on the total change in e, attribute-basing will be unable to equalize these shadow prices, whereas trading would.

Second, the derivation above assumed that a was exogenous, but our empirical evidence below makes clear that this is not so. When a is endogenous, then the attribute-based schedule will induce distortions in the choice of a as described in our baseline model. These costs will be weighed against benefits from marginal cost equalization. It is important to note also that distortion in a is unavoidable in an attribute-based system, so long as a is a choice made by market participants rather than an immutable characteristic. There is no way to design an attribute-based policy to achieve the first-best, even if all marginal costs line up and the R^2 is one, unless the attribute slope is zero.

These deviations from the first-best represent the costs of using attribute-basing in lieu of an efficient policy that equalizes marginal costs across all sources of the externality, as would occur either in a regulation with full trading or in a Pigouvian tax. In the non-trading case, attribute-based regulation can generate benefits from marginal cost equalization, but equalizing marginal costs via attribute-basing will inevitably create distortions in the choice of a, which may in turn distort the choice of e, and will invariably fail to fully equalize marginal costs in real world circumstances. In contrast, trading would equalize marginal costs and create efficiency. Moreover, a uniform *tax* on e (as opposed to a uniform regulation) would achieve the first-best regardless of the structure of and dispersion in costs. Thus, even in markets where trading is difficult or costly, there is an efficient instrument available.

2.5 What explains the use of attribute-basing?

Given the arguments made above, one might ask why attribute-basing exists. One possibility is that policymakers do not fully understand the perverse incentives and inefficiencies of these policies because they had not been previously modeled. A second possibility is that policymakers are forced to design the best possible regulatory policy without trading, due to some political economy forces that constrain the available policy instruments. In that case, our results describe and estimate the costs of operating under those constraints when making policy.

A third possibility regards distribution. If policymakers wish to transfer surplus from the consumers and producers of goods with low a towards those with high a, an attribute-based regulation will achieve that to some degree. For the case of the U.S., it is widely believed that the Detroit Three pushed heavily for attribute-basing because on average they produce heavier, less fuel-efficient vehicles than their competitors, particular Toyota, Honda and Nissan. It is clear that CAFE's attribute-basing makes compliance relatively easier for the U.S. companies. If this was a desired policy goal, then the costs that we document here from distortionary responses and inefficiencies represent the costs of achieving this transfer. Note that heavier vehicles tend to be more expensive (particularly given that cars and light-trucks are regulated separately) and thus are generally purchased by wealthier consumers. As a result, it is likely that wealthier consumers, on average, benefit from the attributing-basing. A distributional justification for attribute-basing is thus more likely to be made on the producer side than the consumer.

A fourth possibility is that policymakers may believe that attribute-basing is necessary to spur innovation. A firm can comply with a tightened regulation by downsizing their products (reduce a so that it is easier to raise e) or by investing in technology (to raise e holding a constant). In a static model, efficiency requires that all margins of adjustment be used until their marginal costs are equalized, but if the technology margin lowers production costs in the future period, then a dynamic model might suggest otherwise. In particular, if there are spillovers from technology, a policymaker might wish to increase the share of the adjustment due to technology beyond that which the market will choose on its own.

In appendix C we explore the possibility that attribute-basing is motivated by a belief that there are spillovers from technological advancement. In this case, the optimal choice of a policy instrument

$t(e - \sigma(a))$ will involve attribute-basing, but this is an artifact of a setup that assumes that both e and a have positive externalities attached to them, due to their relationship with technology. In that setting, attribute-basing provides a second instrument that can affect a Pigouvian tax on a. The optimal slope of the attribute function in this case does not resemble the attribute functions chosen by actual policymakers. Thus, while technological innovation is a plausible reason why some attribute-basing could be desirable, it does not appear to rationalize the policies that we actually see.

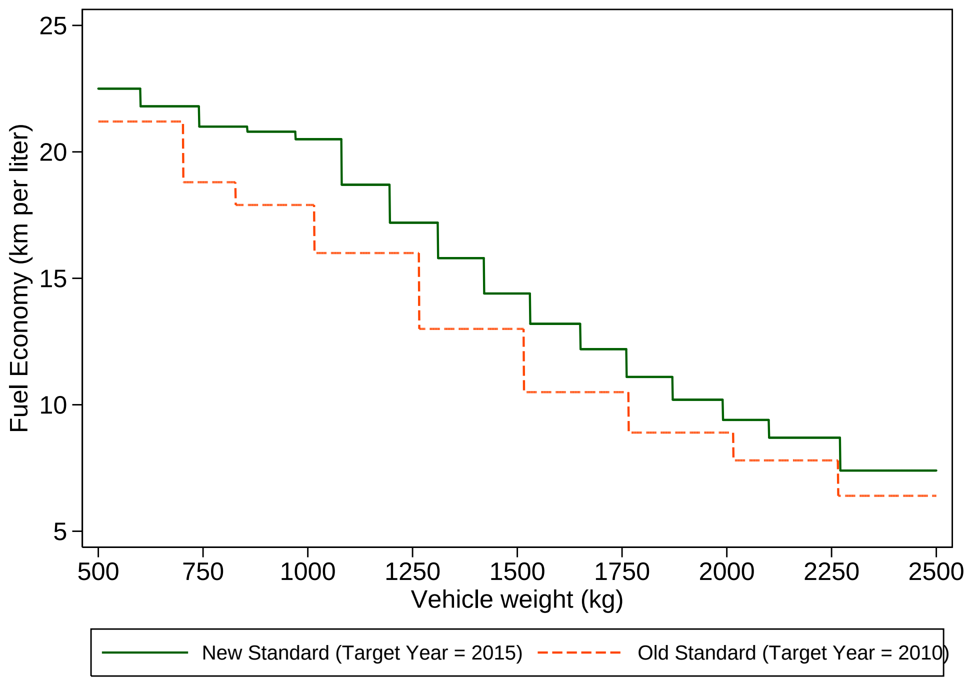
3 Empirical Analysis and Results

In this section, we investigate how the market responds to attributed-based regulation by analyzing the automobile market in Japan. The Japanese context provides several key advantages. First, the Japanese government has been using attribute-based fuel economy regulation since the 1970's. This gives us a long window for analysis and allows us to observe changes in policy that aid identification. Second, the Japanese fuel economy regulations feature notches in the attribute (weight) that determines the fuel economy target. These notches provide substantial variation in regulatory incentives and allow us to use empirical methods similar to those recently developed in the literature on bunching in nonlinear income tax schedules (Saez 1999 2010; Chetty, Friedman, Olsen, and Pistaferri 2011; Kleven and Waseem 2013). In contrast, the U.S. instituted attribute-based regulation for automobiles only in 2012, which leaves us little data for analysis, and its attribute-based target function is smooth, making identification more challenging.

3.1 Fuel Economy Standards in Japan

The Japanese government introduced its first attribute-based fuel economy standard in 1979. Since then, there have been four different schedules. Our data, which start in 2001, come from years spanning the two most recent regimes. We summarize those two policies in Figure 4, which plots the fuel economy target for vehicles as a function of their weight. The downward slope in the figure shows that the fuel economy required of each vehicle is a decreasing function of its weight, so that heavier vehicles face a less stringent regulation. Furthermore, the regulation is notched. The declining function is not smooth, but rather is discontinuous at many weight cutoff points. Thus,

Figure 4: Fuel Economy Standard in Japan



Note: The dashed line shows the fuel economy standard in Japan until 2010. The solid line shows the new fuel economy standards whose target year is 2015.

for vehicles that have a weight near a cutoff point, marginal changes in weight lead to discrete changes in the fuel economy standard. This attribute-basing implicitly subsidizes vehicle weight, which we expect will lead to a distortion by leading to an increase in weight among firms attempting to meet the standard.

Automakers have two incentives to meet these fuel economy standards. First, they are obligated to have their sales-weighted average fuel economy exceed the minimum standard.[13] Second, in some years there are subsidies and tax exemptions that apply to individual cars if their fuel economy exceeds the standard for their weight class by a certain percentage. In 2009, the government introduced these subsidies as part of an economic stimulus package. Since then, the government has changed the eligibility requirements and the size of the subsidies.

An important difference between the fuel economy regulation and the subsidy policy is that the

[13]Technically, this obligation extends to each weight segment separately, but firms are allowed to apply excess credits from one weight category to offset a deficit in another. Thus, in the end, the policy is functionally similar to the U.S. CAFE program, where there is one fleetwide requirement. Automakers have to pay fines if they cannot meet the standard by the target year of the regulation.

later provides an incentive for an individual car to meet the fuel economy standard (as opposed to the fleet), and it thus creates a subsidy system that is notched in both fuel economy and weight. By contrast, the fuel economy regulation does not provide a direct incentive for an individual car to meet the target because what automakers have to do is to meet the target at the sales-weighted average. In the current version of this paper, we recognize that automakers have only the first incentive before 2009 and that they have both incentives in 2009 and 2012, but we do not currently exploit the notches in the fuel economy dimension.

3.2 Data

We analyze data from the Japanese Ministry of Land, Infrastructure, Transportation, and Tourism (MLIT) that records fuel economy data for all new vehicles sold in Japan for each year between 2001 and 2013. The record includes each vehicle's model year, model name, manufacturer, engine type, displacement, transmission type, weight, fuel economy, fuel economy target, estimated carbon dioxide emissions, number of passengers, wheel drive type, and devices used for improving fuel economy. The Ministry data does not include sales volumes. Table 5 presents summary statistics of these data. There are between 1,100 and 1,700 different vehicle configurations sold in the Japanese automobile market each year. This includes both domestic and imported cars. There is a positive trend in weight and displacement over time.

Although most of our analysis uses cross-sectional data, we do conduct panel data analysis in Section 3.5. To do so, we match data between different years in the following way. In our data set, each observation has a product identifier (ID). The product ID is a narrower definition than model name. For example, a Honda Civic may have several product IDs in the same year because there are Civics with different transmissions, displacements and drive types. We first match on product ID across years, which is often, but not always, constant over time. If automakers change the product ID between years, we match by using model name, displacement, drive type (e.g. four-wheel drive), and transmission (manual or automatic). That is, we consider two cars sold in two different years to be the same if they have matching IDs, or if they have the same model name, displacement, drive type, and transmission. We have experimented with matching on differing sets of characteristics and found that our key results are robust to several alternative criteria.

Figure 5: Summary Statistics

Year	N	Fuel Economy (km/liter)		Vehicle weight (kg)		Displacement (liter)		CO2 (g-CO2/km)	
2001	1441	13.53	(4.58)	1241.15	(356.63)	1.84	(0.98)	195.40	(66.72)
2002	1375	13.35	(4.33)	1263.52	(347.00)	1.86	(0.97)	196.72	(66.26)
2003	1178	13.78	(4.53)	1257.15	(356.28)	1.85	(1.03)	191.88	(68.08)
2004	1558	14.20	(4.78)	1255.37	(364.69)	1.82	(1.03)	184.33	(66.67)
2005	1224	13.30	(4.66)	1324.81	(380.62)	2.00	(1.13)	198.14	(71.62)
2006	1286	13.08	(4.59)	1356.56	(391.13)	2.08	(1.17)	201.78	(72.67)
2007	1298	13.24	(4.78)	1369.41	(399.45)	2.09	(1.22)	200.35	(75.07)
2008	1169	13.38	(4.82)	1390.09	(405.77)	2.14	(1.29)	198.58	(76.27)
2009	1264	13.49	(4.93)	1396.40	(413.76)	2.15	(1.30)	197.73	(76.67)
2010	1300	13.50	(5.04)	1428.27	(438.06)	2.21	(1.30)	198.32	(77.34)
2011	1391	13.95	(5.06)	1437.21	(426.23)	2.19	(1.28)	190.15	(71.60)
2012	1541	14.50	(5.21)	1446.50	(411.87)	2.16	(1.24)	182.05	(67.26)
2013	1706	14.43	(5.40)	1476.79	(400.31)	2.24	(1.24)	183.67	(67.37)

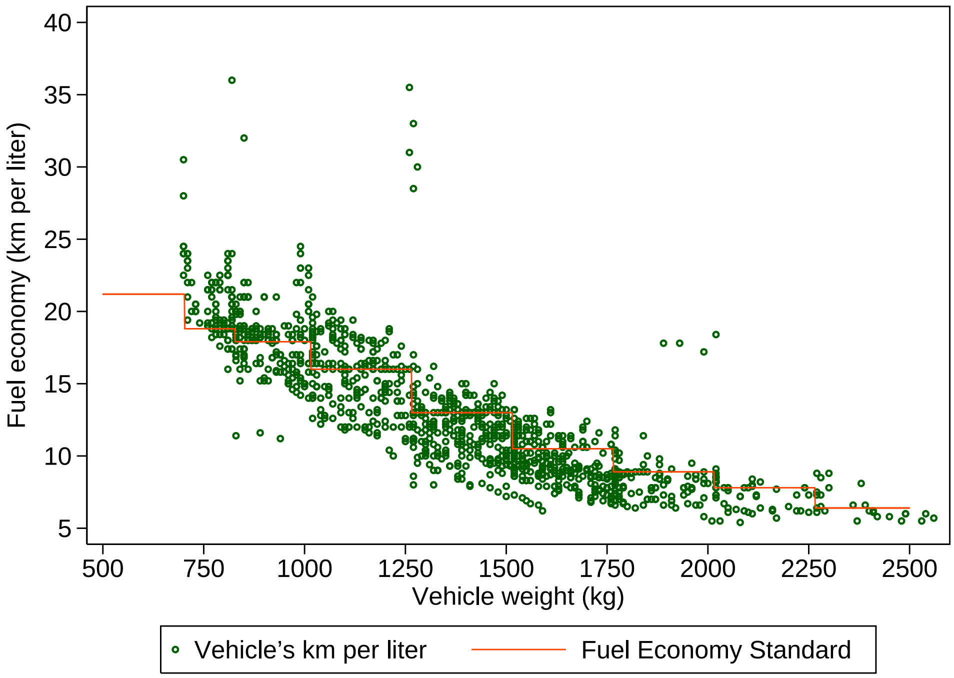
Note: This table shows the number of observations, means and standard deviations of variables by year. Data are not sales weighted.

3.3 Excess Bunching at Notches in Fuel Economy Standard

Our theoretical model emphasize that, in many circumstances, the key empirical statistic that determines the welfare properties of attribute-based standards is the degree to which attribute-basing causes distortions in the choice of the attribute (e.g., weight). For notched policies, distortions in weight should be limited to those vehicles that bunch at a weight notch (see Appendix B for details). We thus begin our analysis by looking at the distribution of the raw data to look for evidence of bunching. Figure 6 plots the fuel economy schedule for 2006, along with a scatter plot of each vehicle's fuel economy and weight in that year. We use 2006 as an example; the pictures for all other years reveal similar patterns.

The diagram reveals three things. First, as expected, there is a negative relationship between fuel economy and vehicle weight. Second, we observe "hugging" of many observations on the fuel economy standard schedule. That is, many vehicles have exactly the same fuel economy as the fuel economy standard, or just slightly higher fuel economy than the standard. Finally, and most prominently, there are excess numbers of observations at the notch points in vehicle weight, which

Figure 6: Scatter Plot of Vehicle Weight (kg) against Fuel Economy (km/liter) in 2006



Note: This figure plots each vehicle's fuel economy (km per liter) in 2006 against the vehicle's weight. Each dot represents one observation.

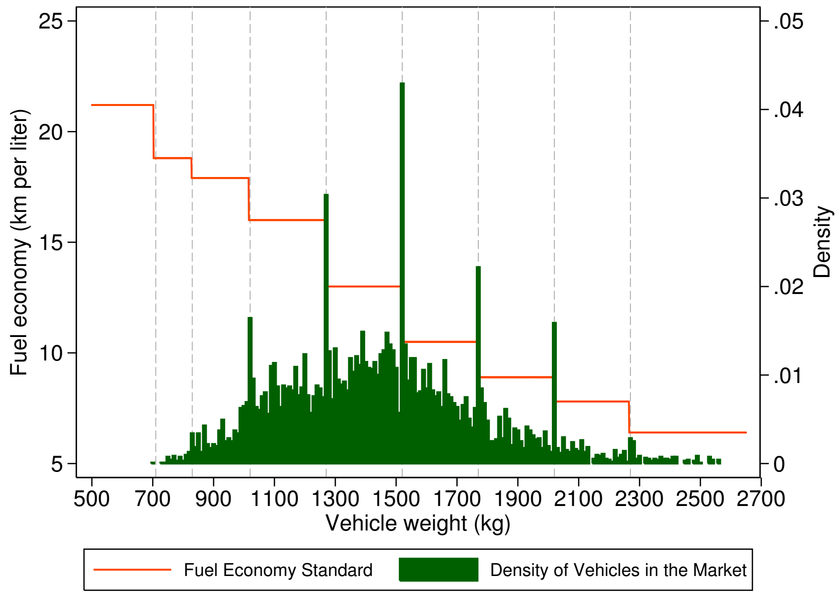
occur at 710, 830, 1020, 1270, 1520, 1770, 2020, and 2270 kg.

To see the excess numbers of observations at the weight notches more clearly, we next provide histograms of vehicle weight superimposed on the notched schedules. Panel A of Figure 7 is the histogram of vehicles sold between 2001 and 2008. The bin size of the histogram is 10 kg. That is, each bar represents the density of the number of observations for each 10 kg bin. In this period, all vehicles had an old fuel economy standard schedule, which is presented in the same figure. There is visually clear excess bunching at each notch of the fuel economy standard schedule. The magnitude of the bunching is substantial; there are more than double the number of observations at each notch point, as compared to the surrounding weight categories.

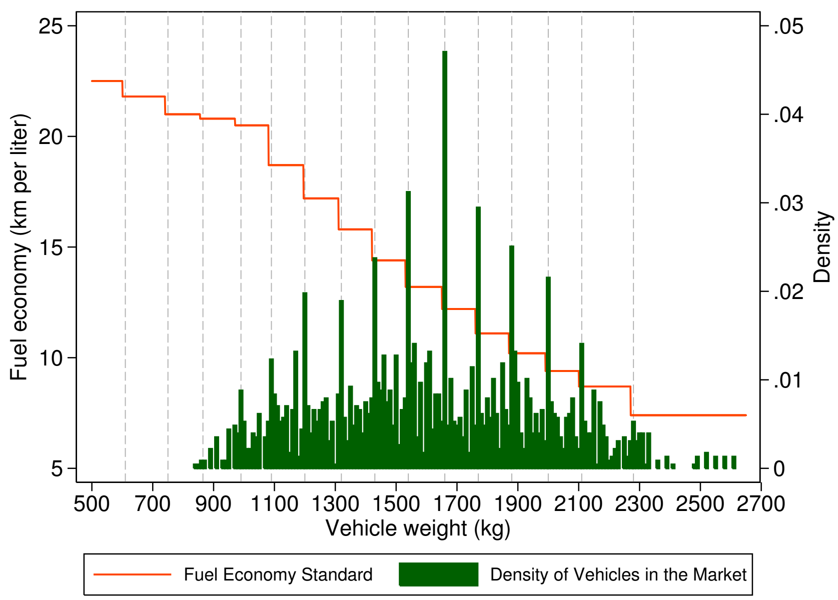
Panel B of Figure 7 shows the corresponding figure for data taken from years when the new standard was in effect. By comparing Panel A and B, one can see that the mass points shifted precisely in accordance with the change in the locations of the notch points. In Figure 8 we present analogous results for "kei-cars", which are very small vehicles that are required to have engine

Figure 7: Fuel Economy Standard and Histogram of Vehicles

Panel A. Years 2001 to 2008 (Old Fuel Economy Standard Schedule)



Panel B. Years 2009 to 2013 (New Fuel Economy Standard Schedule)



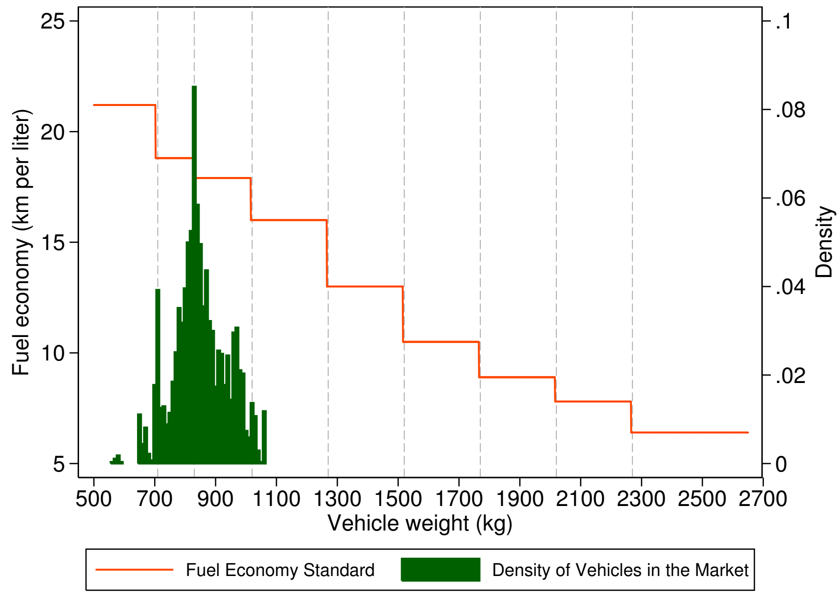
Note: Panel A shows the histogram of vehicles from 2001 to 2008, where all vehicles had the old fuel economy standard. Panel B shows the histogram of vehicles from 2009 to 2013, in which the new fuel economy standard was introduced.

displacements below 660 cc (.66 L).[14] Although kei-cars must comply with the same fuel economy

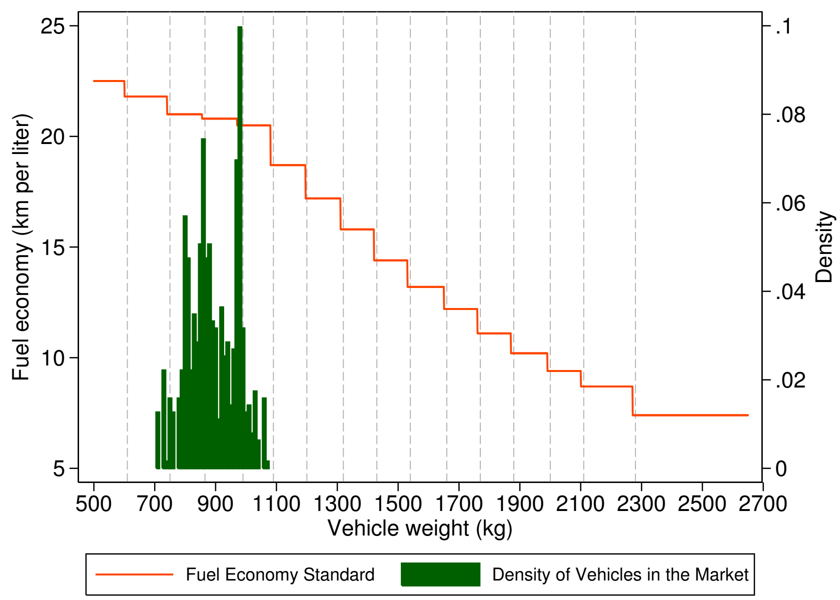
[14]For comparison, no car sold in the U.S. in 2010 would qualify as a kei-car. The two-seat Smart Car has the

Figure 8: Fuel Economy Standard and Histogram of Vehicles: Kei-Cars (small cars)

Panel A. Years 2001 to 2008 (Old Fuel Economy Standard Schedule)



Panel B. Years 2009 to 2013 (New Fuel Economy Standard Schedule)



Note: "Kei-car" is is a Japanese category of small vehicles; the displacement of kei-cars have to be less than 660 cc. Most kei-cars are not exported to other countries. Panel A shows the histogram of vehicles from 2001 to 2008, where all vehicles had the old fuel economy standard. Panel B shows the histogram of vehicles from 2009 to 2013, in which the new fuel economy standard was introduced.

smallest displacement of any car in the U.S. that year, at 100 cc.

regulations as all other cars, we present our results for them separately because they occupy a unique market segment, have different tax and insurance regulations, and are generally viewed as a distinct product category by Japanese consumers. The figure shows that the weight distribution of kei-cars also demonstrates bunching at notch points and appears responsive to changes in the location of notches.

In sum, the raw data show strong evidence that the market responded to the attribute-based regulation by changing vehicle weight. Our theory suggests that, on its face, this is evidence of distortion. In the next section, we use an econometric methods to estimate the magnitude of this excess bunching and discuss its implications for welfare.

3.4 Estimation of Excess Bunching at Notches

Econometric estimation of excess bunching in kinked or notched schedules is relatively new in the economics literature. Saez (1999) and Saez (2010) estimate the income elasticity of taxpayers in the U.S. with respect to income tax rates and EITC schedules by examining excess bunching around kinks in the U.S. personal income tax schedule. Similarly, Chetty, Friedman, Olsen, and Pistaferri (2011) estimate the income elasticity of taxpayers in Denmark with respect to income tax rates by examining the excess bunching in the kinked tax schedules there. In Pakistan, the income tax schedule has notches instead of kinks. That is, the *average* income tax rate is piecewise linear. Kleven and Waseem (2013) uses a method similar to Chetty, Friedman, Olsen, and Pistaferri (2011) to estimate the elasticity of income with respect to income tax rates using bunching around these notches. Our approach is closely related to these papers, although our application is a fuel economy regulation, not an income tax.

To estimate the magnitude of the excess bunching, our first step is to estimate the counterfactual distribution as if there were no bunching at the notch points, which parallels the procedure in Chetty, Friedman, Olsen, and Pistaferri (2011). We start by grouping vehicles into small weight bins (10 kg bins in the application below). For bin j, we denote the number of vehicles in that bin by c_j and the vehicle weight by w_j. For notches $k = 1, ..., K$, we create dummy variables d_k that equal one if j is at notch k. We then fit a polynomial of order S to the bin counts in the empirical distribution, *excluding* the data at the notches, by estimating a regression:[15]

$$c_j = \sum_{s=0}^{S} \beta_s^0 \cdot (w_j)^s + \sum_{k=1}^{K} \gamma_k^0 \cdot d_k + \varepsilon_j, \tag{17}$$

where β_s^0 is an initial estimate for the polynomial fit, and γ_k^0 is an initial estimate for a bin fixed effect for notch k. (We refer to these as initial estimates because we will adjust them in a subsequent step.) By including a dummy for each notch, the polynomial is estimated without considering the data at the notches. We define an initial estimate of the counterfactual distribution as the predicted values from this regression omitting the contribution of the notch dummies: $\hat{c}_j^0 = \sum_{s=0}^{q} \hat{\beta}_s^0 \cdot (w_j)^s$. The excess number of vehicles that locate at the notch relative to this counterfactual density is $\hat{B}_k^0 = c_k - \hat{c}_k^0 = \hat{d}_k^0$.

This simple calculation overestimates B_k because it does not account for the fact that the additional vehicles at the notch come from elsewhere in the distribution. That is, this measure does not satisfy the constraint that the area under the counterfactual distribution must equal the area under the empirical distribution. To account for this problem, we must shift the counterfactual distribution upward until it satisfies this integration constraint.

The appropriate way to shift the counterfactual distribution depends on where the excess bunching comes from. Our theory indicates that attribute-based fuel economy regulation provides incentives to *increase* vehicle weight—that is, excess bunching should come from the "left". We assume that this is the case. We also make the conservative assumption that the bunching observed at a given notch comes only from the adjacent weight bin, which limits the maximum increase in weight.[16] That is, the bunching at notch k comes from bins $j \in (k-1, k)$.[17] In practice, automakers may increase the weight of a vehicle so that it moves more than one weight category. In that case, our procedure will underestimate weight distortions. In this sense, our procedure provides a lower bound on weight manipulation.

[15] We use $S = 7$ for our empirical estimation below. Our estimates are not sensitive to the choice of S for the range in $S \in [3, 11]$.

[16] This restriction is consistent with panel data estimates of weight changes in response to the policy change that we derive below.

[17] For notch $k = 1$ (the first notch point), we use the lowest weight in the data as the minimum weight for this range. Note that this approach may underestimate the change in weight, because the minimum weight in the counterfactual distribution can be lower than the minimum weight in the observed distribution if the attribute-based regulation shifted the minimum weight upward. We want to use this approach to keep our estimate of the change in weight conservative.

In addition, estimation requires that we make some parametric assumption about the distribution of bunching. We make two such assumptions, the first of which follows Chetty, Friedman, Olsen, and Pistaferri (2011), who shift the affected part of the counterfactual distribution uniformly to satisfy the integration constraint. In this approach, we assume that the bunching comes uniformly from the range of $j \in (k-1, k)$. We define the counterfactual distribution $\hat{c}_j = \sum_{s=0}^{q} \hat{\beta}_s \cdot (w_j)^s$ as the fitted values from the regression:

$$c_j + \sum_{k=1}^{K} \alpha_{kj} \cdot \hat{B}_k = \sum_{s=0}^{S} \beta_s \cdot (w_j)^s + \sum_{k=1}^{K} \gamma_k \cdot d_k + \varepsilon_j, \tag{18}$$

where $\hat{B}_k = c_k - \hat{c}_k = \hat{d}_k$ is the excess number of vehicles at the notch implied by this counterfactual. The left hand side of this equation implies that we shift c_j by $\sum_{k=1}^{K} \alpha_{kj} \cdot \hat{B}_k$ to satisfy the integration constraint. The uniform assumption implies that we assign $\alpha_{kj} = \frac{c_j}{\sum_{j \in (k-1,k)} c_j}$ for j $\in (k-1, k)$ and $= 0$ for j $\notin (k-1, k)$. Because $\hat{B}_k$ is a function of $\tilde{\beta}_s$, the dependent variable in this regression depends on the estimates of $\tilde{\beta}_s$. We therefore estimate this regression by iteration, recomputing $\hat{B}_k$ using the estimated $\tilde{\beta}_s$ until we reach a fixed point. The bootstrapped standard errors that we describe below adjust for this iterative estimation procedure.

Note that this approach with the uniform assumption may underestimate or overestimate Δw if the bunching comes disproportionally from the range of $j \in (k-1, k)$. For example, if most bunching comes from the bins near k rather than the bins near $k-1$, the uniform assumption approach overestimates Δw. This may not be a substantial issue for our estimation, because the empirical distribution in Figure 7 suggests that there are no clear holes in the distribution. There should be clear holes if the origins of the excess bunching are substantially disproportional to the distribution. However, we prefer an approach that does not impose the uniform assumption. To do that, we propose an approach that defines α_j based on the empirical distribution of vehicles relative to the counterfactual distribution. We define the ratio between the counterfactual and observed distributions by $\theta_j = \hat{c}_j / c_j$ for j $\in (k-1, k)$ and $= 0$ for j $\notin (k-1, k)$. Then, we define $\alpha_{kj} = \frac{\theta_j}{\sum_{j \in (k-1,k)} \theta_j}$. By this approach, α_{kj} is obtained from the relative ratio between the counterfactual and observed distributions. We use this approach for our main estimate and also report estimates from the uniform assumption approach as well.

In addition to $\hat{B}_k$ (the excess number of vehicles at notch k), we provide two more estimates that are relevant to our welfare calculation. The first estimate is the proportional excess bunching, which is defined by $\hat{b} = c_k/\hat{c}_k$. This estimate shows the proportional size of the bunching relative to the height of the counterfactual distribution at k. The second estimate is the average changes in weight (kg) for vehicles at notch k, defined by the weighted average of the estimated change in weight: $E[\Delta w_k] = \frac{\sum_{j\in(k-1,k)} (w_k - w_j)\cdot(\hat{c}_j - c_j)}{\sum_{j\in(k-1,k)} (\hat{c}_j - c_j)}$.

We calculate standard errors using a parametric bootstrap procedure used by Chetty, Friedman, Olsen, and Pistaferri (2011) and Kleven and Waseem (2013). We draw from the estimated vector of errors ϵ_j in (18) with replacement to generate a new set of counts and apply the technique above to calculate new estimates. We define the standard errors as the standard deviation of the distribution of these estimates. Since we observe the exact population distribution of vehicles in the Japanese automobile market, this standard error reflects error due to misspecification of the polynomial for the counterfactual distribution rather than sampling error.

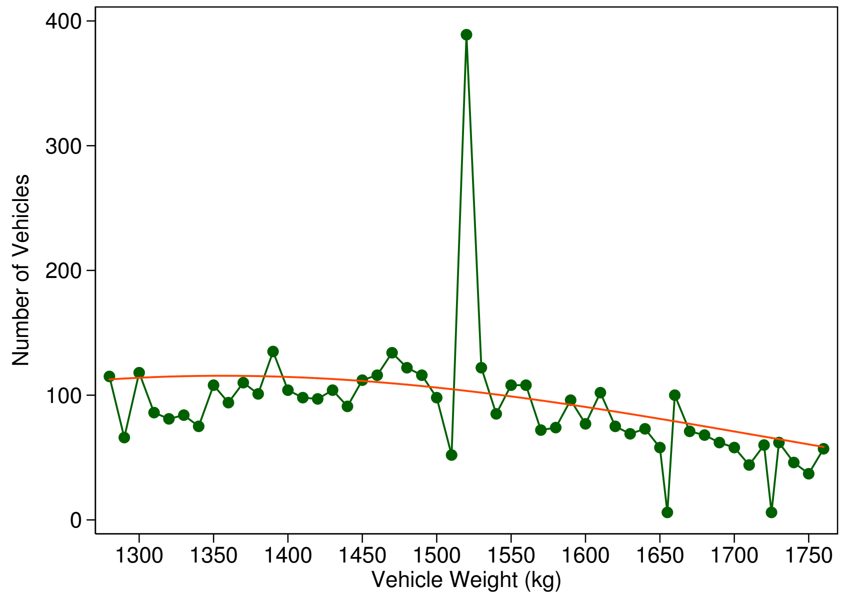
Figure 9 shows our estimation procedure graphically for two notch points. In Panel A, we plot the actual distribution and estimated counterfactual distribution at the 1520 kg notch point. Graphically, our estimate of excess bunching is the difference in height between the actual and counterfactual distribution at the notch point. The estimate and standard error of the excess number of vehicles B is 285.27 (3.75). That is, there are 285 excess vehicles at this notch compared to the counterfactual distribution. The proportional bunching $b = 3.75$ (0.21), which means that the observed distribution has 3.75 times more observations than the counterfactual distribution at the notch. Finally, the average weight increase $E[\Delta w]$ is 114.97 (0.22) kg for affected cars. Similarly, we illustrate our estimation result at the 2020 kg notch point. At this notch, $B = 127.07$ (9.04), $b = 8.51$ (1.55), and $E[\Delta w]$ =120.77 (0.15).

Table 1 presents our estimation results for all notches for the data between 2001 and 2008 (the old fuel economy standard). To see the automakers' incentives at each notch, column 2 shows the stringency of the fuel economy standard (km/liter) below and above the notch (higher km/liter numbers imply more stringent standards). Columns 3 to 5 report our main estimates based on the approach described above. First, we find statistically significant excess bunching at all notches except for the 1020 kg notch for kei-cars, where the estimate is noisy because there is not a large number of kei-cars in this weight range. Second, we find substantial heterogeneity in the estimates

Figure 9: Graphical Illustration of Estimation of Excess Bunching at Each Notch Point

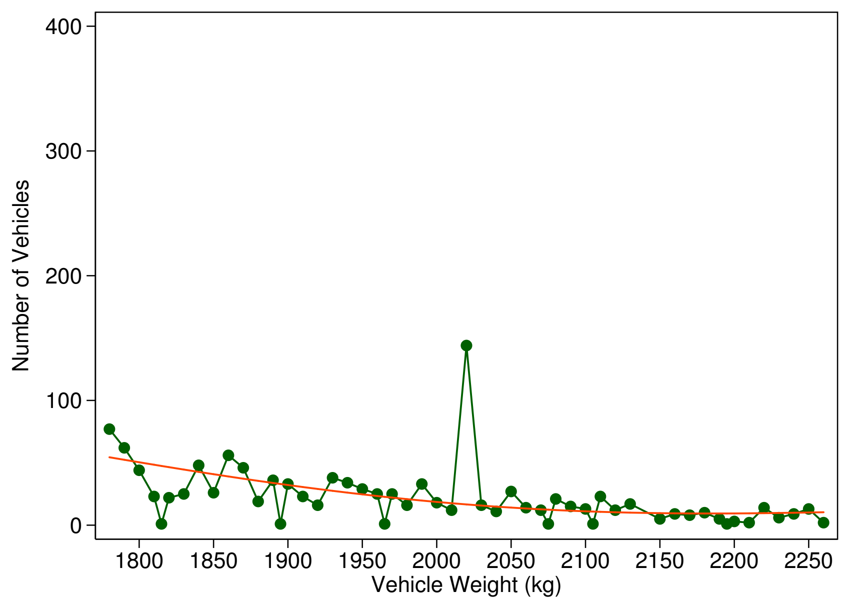
Panel A. Notch at 1520 kg

B = 285.27 (3.75), b = 3.75 (0.21), $E[\Delta w]$ =114.97 (0.22)



Panel B. Notch at 2020 kg

B = 127.07 (9.04), b = 8.51 (1.55), $E[\Delta w]$ =120.77 (0.15)



Note: This figure graphically shows the estimation in equation (18). The figure also lists the estimates of B (excess bunching), b (proportional excess bunching), and $E[\Delta w]$ (the average weight increase). See the main text for details on these estimates.

Table 1: Excess Bunching and Weight Increases at Each Notch: Old Fuel Economy Standard

Notch Point	Fuel Economy Standard below & above the Notch (km/liter)	Main Estimates: Excess Bunching (# of cars)	Main Estimates: Excess Bunching (ratio)	Main Estimates: E[Δweight] (kg)	Uniform Assumption: Excess Bunching (# of cars)	Uniform Assumption: Excess Bunching (ratio)	Uniform Assumption: E[Δweight] (kg)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
830 kg	18.8	16.46	2.13	51.57	16.73	2.17	55.00
	17.9	(7.91)	(0.49)	(3.21)	(7.28)	(0.44)	N.A.
1020 kg	17.9	87.18	2.41	103.77	87.02	2.40	95.00
	16	(8.05)	(0.16)	(0.49)	(7.48)	(0.13)	N.A.
1270 kg	16	163.48	2.47	146.89	163.33	2.46	125.00
	13	(7.92)	(0.11)	(0.62)	(7.33)	(0.08)	N.A.
1520 kg	13	285.27	3.75	114.97	285.41	3.76	125.00
	10.5	(8.21)	(0.21)	(0.22)	(7.52)	(0.15)	N.A.
1770 kg	10.5	143.93	3.52	129.44	144.25	3.54	125.00
	8.9	(8.93)	(0.30)	(0.57)	(8.13)	(0.24)	N.A.
2020 kg	8.9	127.07	8.51	120.77	127.24	8.59	125.00
	7.8	(9.04)	(1.55)	(0.15)	(8.28)	(1.43)	N.A.
2270 kg	7.8	15.67	2.52	137.86	15.52	2.48	125.00
	6.4	(6.40)	(0.66)	(4.48)	(5.95)	(0.64)	N.A.
Kei-Cars							
710 kg	21.2	60.53	2.33	72.57	59.44	2.28	75.00
	18.8	(15.54)	(0.30)	(0.57)	(13.37)	(0.25)	N.A.
830 kg	18.8	118.36	2.06	39.79	120.15	2.09	60.00
	17.9	(15.99)	(0.14)	(0.09)	(12.77)	(0.11)	N.A.
1020 kg	17.9	21.15	2.33	92.63	19.30	2.09	95.00
	16	(9.48)	(2.30)	(1.08)	(7.51)	(2.08)	N.A.

Note: This table shows the regression result in equation 18. Bootstrapped standard errors are in the parentheses.

between the notches. The proportional excess bunching b ranges between 2.1 to 8.5, depending on the notches. The estimated weight increases $E[\Delta w]$ range between 40 kg to 93 kg for kei-cars and 52 kg to 147 kg for other cars. For most cars, this is around 10% increases in weight, which is substantial. Third, our estimates based on the uniform assumption provide similar estimates for the excess number of vehicles B and the proportional excess bunching b. Our estimates for B

Table 2: Excess Bunching and Weight Increases at Each Notch: New Fuel Economy Standard

Notch Point	Fuel Economy Standard below & above the Notch (km/liter)	Main Estimates: Excess Bunching (# of cars)	Main Estimates: Excess Bunching (ratio)	Main Estimates: E[Δweight] (kg)	Uniform Assumption: Excess Bunching (# of cars)	Uniform Assumption: Excess Bunching (ratio)	Uniform Assumption: E[Δweight] (kg)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
980 kg	20.8	5.37	1.60	50.13	5.35	1.60	40.00
	20.5	(5.98)	(0.50)	(1.25)	(5.17)	(0.44)	N.A.
1090 kg	20.5	14.34	2.05	51.58	14.33	2.05	55.00
	18.7	(5.47)	(0.34)	(1.31)	(4.73)	(0.31)	N.A.
1200 kg	18.7	26.56	2.44	61.40	26.58	2.44	55.00
	17.2	(4.85)	(0.24)	(0.26)	(4.19)	(0.24)	N.A.
1320 kg	17.2	20.20	1.89	64.38	20.24	1.89	60.00
	15.8	(4.37)	(0.16)	(0.45)	(3.78)	(0.16)	N.A.
1430 kg	15.8	28.54	2.12	52.18	28.57	2.12	55.00
	14.4	(4.26)	(0.15)	(0.28)	(3.68)	(0.16)	N.A.
1540 kg	14.4	44.44	2.67	61.76	44.45	2.67	55.00
	13.2	(4.45)	(0.18)	(0.10)	(3.84)	(0.21)	N.A.
1660 kg	13.2	81.08	4.13	65.17	81.07	4.13	60.00
	12.2	(4.93)	(0.33)	(0.28)	(4.24)	(0.39)	N.A.
1770 kg	12.2	43.24	2.82	53.86	43.19	2.81	55.00
	11.1	(5.44)	(0.25)	(0.58)	(4.68)	(0.26)	N.A.
1880 kg	11.1	36.60	2.79	52.56	36.52	2.78	55.00
	10.2	(5.86)	(0.29)	(0.49)	(5.04)	(0.30)	N.A.
2000 kg	10.2	33.16	3.09	68.93	33.08	3.08	60.00
	9.4	(5.99)	(0.38)	(0.26)	(5.16)	(0.39)	N.A.
2110 kg	9.4	20.62	2.81	58.08	20.56	2.80	55.00
	8.7	(5.66)	(0.44)	(0.60)	(4.89)	(0.42)	N.A.
2280 kg	8.7	6.66	2.25	91.89	6.65	2.24	88.00
	7.4	(4.09)	(1.13)	(2.21)	(3.57)	(0.87)	N.A.
Kei-Cars							
860 kg	21	18.72	1.66	46.93	18.66	1.66	55.00
	20.8	(4.52)	(0.14)	(0.77)	(4.04)	(0.13)	N.A.
980 kg	20.8	44.82	3.47	55.80	45.04	3.51	60.00
	20.5	(2.31)	(0.40)	(0.97)	(1.80)	(0.54)	N.A.

Note: This table shows the regression result in equation 18. Bootstrapped standard errors are in the parentheses.

and b are not sensitive to this assumption because the excess bunching is so substantial compared to the counterfactual distribution that the way to reach the integration constraint between the observed and counterfactual distributions do not matter much to these estimates. However, we find slight differences in $E[\Delta w]$ between our two methods. With the uniform assumption, $E[\Delta w]$ equals half the length of the immediate left weight category by assumption and therefore we have no standard errors for them. In contrast, our main estimates do not impose the uniform assumption. Our results show that the uniform assumption provides slightly underestimated or overestimated $E[\Delta w]$, depending on the notches. However, the differences are not that large.

Likewise, Table 2 presents our estimation results for all notches for the data between 2009 and 2013 (the new fuel economy standard). Note that the new fuel economy standard has a larger numbers of notches and the length of each weight category is smaller than the ones in the old standard. It means that we expect lower estimates for $E[\Delta w]$, because our conservative estimates assume that automakers may increase their vehicle weight but not by more than the amount to reach the immediate upper weight category. First, we find statistical significant excess bunching at all notches except for the 980 kg notch and 2280 kg notch for normal cars. At these two notches, the estimates indicate that there is positive excess bunching, but the estimates are noisy because there are not a large number of vehicles in this weight range. Second, similar to the estimates in the old standard, we find substantial heterogeneity in the estimates between the notches. The proportional excess bunching b ranges between 2.1 and 4.1 for normal cars and 1.7 and 3.5 for kei-cars, depending on the notches. The estimated weight increases $E[\Delta w]$ range between 50 kg to 92 kg for normal cars and 47 kg to 56 kg for kei-cars. Finally, similar to our results for the old standard, the method with the uniform assumption provides similar estimates to our main estimates.

Overall, the results in this section provide evidence that automakers significantly respond to the attribute-based fuel economy regulation by changing their vehicle weight. By using the bunching estimation method based on the substantial cross-sectional variation, we show that given the assumption that bunching comes from the left (lower weight) of each notch, our estimates of excess bunching imply that automakers increase vehicle weight between 50 kg and 147 kg to reach less stringent fuel economy regulations. In addition to the cross-sectional variation, the Japanese policy change provides useful panel variation in the regulation. In the next section, we exploit our panel data and a policy change in fuel economy regulation.

3.5 Panel Data Analysis: Exploiting the Policy Change

The previous section provided visually clear and statistically significant evidence of bunching at the notch points in the weight dimension using cross-sectional data. In this section, we exploit the policy change in the Japanese fuel economy standard that changed the location of the notched schedules to verify our results using panel techniques.

The two policy regimes are shown in Figure 4. Before 2008, automakers faced the old fuel economy standard represented by the dashed line. In 2009, the new standard was introduced. The years between 2009 and 2012 were a transition period, during which time automakers were allowed to choose which schedule they wished to use. In 2013, all automakers were required to be on the new standard. We focus our comparison on 2008 and 2013 because the automakers did not have a choice about the schedule in those years. For these two years, we use our panel data, which links the same car type over time, to see how vehicle weight and fuel economy changed.

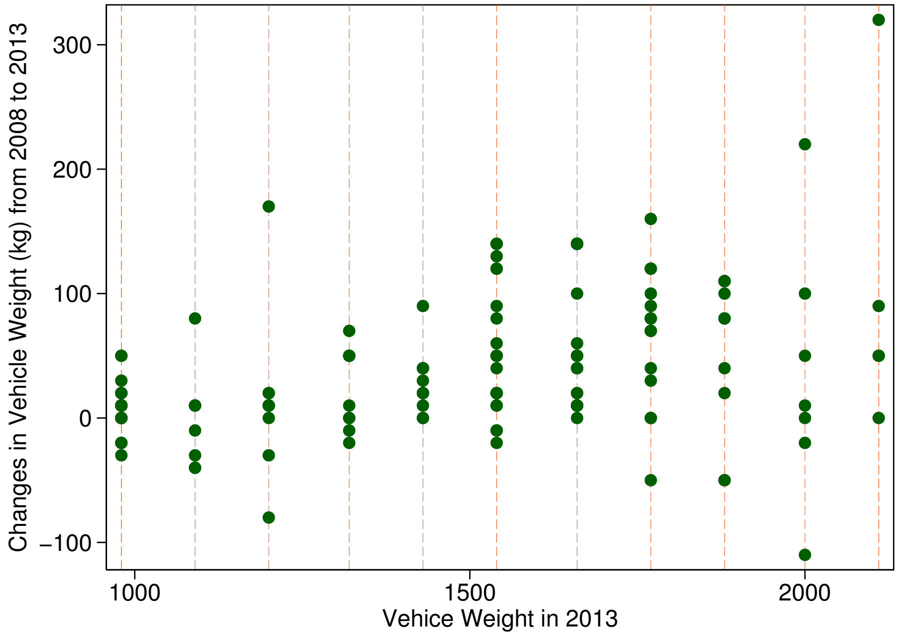
Focusing on the comparison between 2008 and 2013 makes the analysis cleaner, although it comes at a cost. In five years, automakers often change their model names completely, introduce new models, and terminate the production of others. As a result, We are able to match only about half of the vehicles between 2008 and 2013. All results below are based on this matched sample.

We first use our panel data to verify that vehicles that bunch at notches do so by increasing their weight, as suggested by theory, rather than decreasing it. Panel B of Figure 7 shows that many cars had weights within 10 kg of the notches under the 2013 standard. In Figure 10, we take the sample of those vehicles and plot their change in weight between 2008 and 2013 against their final weight in 2013. The horizontal axis shows the vehicle weight in 2013, the vertical axis shows the change in weight from 2008 to 2013, and the dashed vertical lines show where there are weight notches. For most vehicles in the diagram, the change in weight is positive, which suggests that bunching comes from weight increases.

The weight increase in Figure 10 could reflect a secular trend in weight. To account for this, we use the vehicles that do not bunch in 2013 as a control group.[18] Figure 11 shows the kernel density of the changes in weight from 2008 to 2013 for the vehicles that bunched in 2013 and the

[18]We focus on vehicles that have a weight between 960 and 2020 kg, which represents a significant majority of the market, as a control group because light vehicles (those weighing less than 800 kg) appear to have had a significantly different weight trend. Adding the excluded light cars does not change results in our regression analysis, below, in which we include a polynomial that allows the trend to differ vehicles of different weights.

Figure 10: Changes in Vehicle Weight from 2008 to 2013 for Samples that Bunch in 2013



Note: This figure plots the raw data of individual vehicles that bunch in the notch points in the fuel economy standard in 2013. The horizontal axis shows the vehicle weight in 2013, the vertical axis shows the change in weight from 2008 to 2013, and the dashed vertical lines present the notch points.

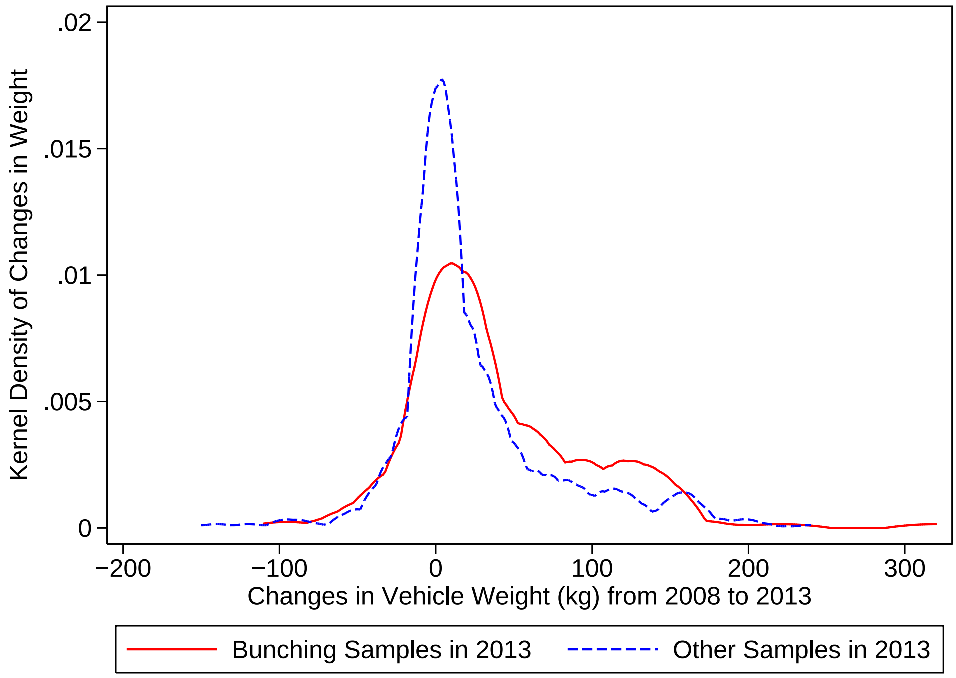
same density for those that did not. The sample of those who bunched is pushed towards the right relative to the control group of non-bunchers. In particular, the distribution indicates that there are many vehicles who bunched at a weight notch after increasing weight by 50 to 150 kg.

To estimate the difference in the change in weight between the bunching and non-bunching vehicles, we estimate the following equation by OLS:

$$\Delta weight_i = \alpha + \beta \cdot 1\{bunching_i\} + f(weight_{i,2013}) + \varepsilon_i, \tag{19}$$

where $\Delta weight_i$ is the change in weight (kg) from 2008 to 2013 for vehicle i and $1\{bunching_i\}$ equals one if vehicle i is located within 10 kg of a weight notch in 2013. Regressing $\Delta weight_i$ on $1\{bunching_i\}$ estimates a difference in mean weight between the two groups. This simple regression potentially produces a bias if the change in weight is systematically larger for lighter or heavier vehicles. In our case, any such bias is likely to be small because the notch points are spread

Figure 11: Kernel Density of Changes in Vehicle Weight from 2008 to 2013



Note: This figure shows the kernel density of the changes in weight from 2008 to 2013 for the bunching samples and other samples in 2013.

throughout the weight distribution, which means that the bunching and non-bunching samples have fairly similar weight on average. Nevertheless, we add a polynomial in $weight_{i,2013}$ to reduce this potential bias.

Table 3 presents regression results. Column 1 shows that the simple mean difference in the change in weight between the bunching samples and others is 13.79 kg. That is, vehicles that bunched increased their weight by 13.79 kg more than other vehicles. In column 2, we control for any linear relationship between the change in weight and final weight by adding a linear control for weight in 2013. As expected, this yields a positive coefficient on weight, which means that heavier vehicles had slightly larger increases in weight, but this has very little impact on our coefficient of interest, as expected. We include the second- and third-order polynomials in columns 3 and 4, which has no significant impact on our coefficient of interest.

This evidence confirms our theoretical prediction that automakers *increase* their vehicle weight in response to the attribute-based regulation notches. We estimate that the mean of this weight

Table 3: Changes in Vehicle Weight from 2008 to 2013: Does Bunching Come from Left or Right?

	(1) ΔWeight	(2) ΔWeight	(3) ΔWeight	(4) ΔWeight
1{Bunching in 2013}	13.79***	13.23***	13.62***	12.48**
	(5.25)	(5.04)	(4.98)	(5.07)
(Weight in 2013)/1000		46.14***	261.34***	-131.18
		(6.83)	(54.79)	(346.21)
(Weight in 2013)/1000)^2			-71.47***	185.45
			(18.06)	(224.48)
((Weight in 2013)/1000)^3				-54.03
				(47.05)
Constant	24.56***	-42.32***	-196.26***	-4.05
	(2.81)	(10.26)	(40.19)	(172.16)
Observations	531	531	531	531
R^2	0.013	0.091	0.118	0.120

Note: This table shows the regression result in equation 19. The dependent variable is the change in vehicle weight (kg) from 2008 to 2011. 1{Bunching in 2013} is a dummy variable or samples that bunch in the fuel economy standard schedule in 2013. Standard errors are in the parentheses. ***, **, and * show 1, 5, and 10% levels of statistical significance from zero.

increase due to the change in the policy is about 13 kg. This estimate is smaller than our estimate in the previous section, which uses cross-sectional variation at the notches. This is not surprising because the two identification strategies use different sizes of variation in fuel economy standards. The cross-sectional variation at the notches is much larger than the panel-variation that we use in this section, because the panel-variation exploits the *changes* in fuel economy standard between the old and new fuel economy standards. We go further in the next section by using the panel variation in incentives created by the policy change to estimate a weight response elasticity.

3.6 Panel Data Analysis: Elasticity of Attributes with Respect to Regulation

Our theory indicates that if a planner were to choose the second-best tax rate on fuel economy, taking as given the slope of the attribute-based target, that the deviation between the second-best tax rate and the Pigouvian benchmark would be a function of the relative elasticities of weight and fuel economy when the policy was tightened. In this section we use panel variation to estimate those elasticities and provide a back of the envelope calculation as to how much the attribute-basing would lower the optimal (second-best) stringency of the policy.

Figure 4 shows that the new regulation increased the fuel economy requirement for vehicles in all weight classes, but the increase was of a different magnitude for vehicles in different weight classes. For example, compare a vehicle that weighed 1420 kg in 2008 to one that weighed 1430 kg in that year. The former vehicle faced an increase in its fuel economy standard from 13 to 15.8 km/liter (a 22% increase), while the latter vehicle faced an increase in its fuel economy standard from 13 to 14.3 km/liter (an 11% increase).[19] In this section, we exploit the discontinuities in the variation in the changes in fuel economy standards to estimate the elasticity of attributes with respect to fuel economy regulation.

We use a regression discontinuity design (RDD) that is similar to the identification strategy in Saez (2003). Saez uses discontinuous variation in income tax rates created by "bracket creep" in the U.S. income tax schedule. In his study, there is no policy change in the income tax schedule in nominal terms. However, inflation is high enough to shift the income tax schedule towards the right in real terms, causing taxpayers to face differing real income tax schedules. In our context, we have changes in 1) the locations of the kink points and 2) the stringency of the fuel economy regulation, which provides powerful variation enabling us to estimate how variation in the stringency of regulation affects attribute changes.

Our estimation is based on the following OLS regression:

$$\Delta weight_i = \alpha + \beta \cdot \Delta standard_i + \gamma \cdot X_i + \varepsilon_i, \tag{20}$$

where $\Delta weight_i$ is the change in weight (kg) from 2008 to 2013 for vehicle i, $\Delta standard_i$ is the change in the fuel economy standard (km/liter) for i, and X_i is a vector of control variables. OLS estimates will be biased because $standard_{it}$ is a function of $weight_{it}$ for $t \in (2008, 2013)$. If there are unobservable shocks in ε_i, they will affect $\Delta standard_i$ as well as $\Delta weight_i$. Because the fuel economy standard is a decreasing step function of weight, we expect that OLS estimation produces a downward bias for β.

To address the endogeneity, we use a policy-induced change in the standard as an instrument, as in Saez (2003); Saez, Slemrod, and Giertz (2012); Ito (Forthcoming). Specifically, we use

[19] Note that this variation is driven by the front-runner in each category, which we treat as plausibly random variation.

Table 4: Elasticity of Attributes (weight) with Respect to Fuel Economy Standards

	ΔWeight(kg)			Δln(Weight)	
	(1)	(2)	(3)	(4)	(5)
ΔFuel Economy Standard	-8.67***	18.70**	18.86***		
	(1.69)	(7.37)	(7.26)		
Δln(Fuel Economy Standard)				0.16***	0.16***
				(0.05)	(0.05)
(Weight in 2008)/1000		479.58***	313.30	0.24***	0.18
		(76.40)	(434.38)	(0.04)	(0.26)
(Weight in 2008)/1000)^2		-146.39***	-36.23	-0.08***	-0.04
		(24.04)	(278.79)	(0.01)	(0.17)
((Weight in 2008)/1000)^3			-23.40		-0.01
			(58.19)		(0.04)
Constant	42.00***	-371.45***	-291.45	-0.18***	-0.14
	(3.51)	(66.05)	(224.59)	(0.03)	(0.13)
Observations	531	531	531	531	531
Estimation	OLS	RD(2SLS)	RD(2SLS)	RD(2SLS)	RD(2SLS)

Note: This table shows the regression result in equation 20. The dependent variable is the change in vehicle weight (kg) from 2008 to 2011 or log of the change in weight. Δ Fuel Economy Standard is the change in fuel economy standard. Standard errors are in the parentheses. ***, **, and * show 1, 5, and 10% levels of statistical significance from zero.

$\Delta standard_i^{PI} = standard_{i,2013}(weight_{i,2008}) - standard_{i,2008}(weight_{i,2008})$. This instrument, which is sometimes called a simulated instrument, computes the predicted change in the standard at a weight level $weight_{i,2008}$. The instrument thus captures the change induced by the policy change for the weight level $weight_{i,2008}$. To be a valid instrument, $weight_{i,2008}$ has to be uncorrelated with ε_i, because the instrument is a function of $weight_{i,2008}$. This condition can be violated if lighter cars and heavier cars in 2008 have different underlying changes in weight. The advantage of our identification strategy is that we can include any smooth controls for $weight_{i,2008}$ in X_i to account for such concerns. Because the instrument's variation in the change in standard is discontinuous in weight, including any smooth controls for $weight_{i,2008}$ does not eliminate our ability to identify the coefficient. In our estimation, we use $\Delta standard_i^{PI}$ as an instrument for $\Delta standard_i$ and include a third-order polynomial of $weight_{i,2008}$ in X_i.[20]

[20]Including yet higher orders of the polynomial does not change our results.

Table 5: Elasticity of Attributes (fuel economy) with Respect to Fuel Economy Standards

	ΔFuel Economy (km/liter)			Δln(Fuel Economy)	
	(1)	(2)	(3)	(4)	(5)
ΔFuel Economy Standard	-0.00	0.22	0.20		
	(0.06)	(0.21)	(0.21)		
Δln(Fuel Economy Standard)				0.22	0.21
				(0.16)	(0.16)
(Weight in 2008)/1000		7.80***	23.07*	0.56***	1.25
		(2.18)	(12.34)	(0.13)	(0.79)
(Weight in 2008)/1000)^2		-2.61***	-12.73	-0.19***	-0.65
		(0.69)	(7.92)	(0.04)	(0.51)
((Weight in 2008)/1000)^3			2.15		0.10
			(1.65)		(0.11)
Constant	0.26**	-5.61***	-12.95**	-0.40***	-0.73*
	(0.12)	(1.89)	(6.38)	(0.10)	(0.39)
Observations	531	531	531	531	531
Estimation	OLS	RD(2SLS)	RD(2SLS)	RD(2SLS)	RD(2SLS)

Note: This table shows the regression result in equation 20 but uses $\Delta FuelEconomy_i$ as the left-hand side variable. $\Delta FuelEconomy_i$ is the change in fuel economy (km/liter) for vehicle i from 2008 to 2013. Standard errors are in the parentheses. ***, **, and * show 1, 5, and 10% levels of statistical significance from zero.

Table 4 presents estimation results. As expected, the OLS estimate in column 1 is lower than our RD estimates in columns 2 and 3. This is likely due to the downward bias stemming from the endogeneity of $\Delta standard_i$. In columns 2 and 3, we show our RD estimates. The first stages are very strong because the policy-induced change in fuel economy standard strongly affects the actual change in standard. After we include the first- and second-order of polynomials in $weight_{i,2008}$, our estimates are not sensitive to the inclusion of higher-orders of polynomials. The estimates imply that a one unit increase in the fuel economy standard (km/liter) results in an increase in vehicle weight of 19 kg. In columns 4 and 5, we use a log-log RD specification to estimate a constant elasticity. The estimate implies that a 1% increase in the fuel economy standard results in a 0.16% increase in vehicle weight. These results provide empirical evidence that automakers change their vehicle weight in response to the change in the stringency of fuel economy regulation.

In Table 5, we estimate equation 20 but use $\Delta FuelEconomy_i$ as the left-hand side variable. $\Delta FuelEconomy_i$ is the change in fuel economy (km/liter) for vehicle i from 2008 to 2013. In this estimation, we test whether increases in the fuel economy standard cause an increase in the fuel economy of vehicles. The estimates indicate that more stringent fuel economy standards result in increases in fuel economy, but the estimates are not statistically significantly different from zero.

We can use equation 7 to translate these elasticity estimates into a policy implication. The formula relates the second-best tax rate (or shadow price) t to the externality ϕ and the elasticities of a and e with respect to the policy, and the elasticity of the target fuel economy rating with respect to weight:

$$t^* = \frac{\phi}{1 - \epsilon_a^\sigma \frac{\epsilon_t^a}{\epsilon_t^e}} = \frac{\phi}{1 - (-.89)\frac{.16}{.21}} = .60\phi.$$

We estimate the elasticity of the policy function by regressing the log of the fuel economy values for each step on the log of the weight, measured at the midpoint, of the steps for the 2008 policy and get an elasticity of -.89. We then plug in our estimates of the elasticities of weight and fuel economy with respect to a tightening of the policy from tables 4 and 5. Together, these estimates imply that the second-best tax rate would be equal to marginal damages divided by 1.68, which is equivalent to saying that the second-best tax rate is 60% of the Pigouvian tax. That is, because weight responses strongly to increases in the standard, the policy-maker would choose a tax rate significantly below marginal damages from the externality in order to balance the distortions from the choice of a against the externality improvement.[21]

In summary, our empirical analysis provides three findings: 1) there is compelling evidence that vehicles bunch at the attribute (weight) notch points in the Japanese fuel economy standard; 2) given the assumption (backed by theory) that bunching comes from the left (i.e., bunching comes from increases in weight), the excess bunching in our cross-sectional variation implies that the average increase in vehicle weight for each notch point is between 40 kg and 147 kg; and 3) our panel data analysis provides supporting evidence for this assumption, as we find that a one unit increase in the fuel economy standard results in an increase in vehicle weight of 19 kg and a one percent increase in fuel economy standard results in an increase in vehicle weight by 0.16%.

[21] Note that our estimate of the fuel economy elasticity is imprecise, so that the standard error around this statistic is large, though even the largest values of the fuel economy elasticity still imply significant deviations from Pigou. For example, if we take the upper value of the elasticity's 95% confidence interval, but continue to use the point estimate for the weight elasticity, the reduction in the second-best tax relative to the Pigouvian benchmark would be 21%.

3.7 Welfare implications of weight manipulation

As emphasized in the theory section, the main welfare loss from the response to attribute based standards in the situation where the attribute itself has an unpriced externality will be the inefficiency from an increase in the attribute. An approximation of the distortion from the policy can therefore be calculated by multiplying the change in the attribute by the externality, as indicated in equation 8. Our cross-sectional estimates suggest that the increase in weight caused by attribute-basing in the Japanese car policy is between 40 and 147 kg. The weighted average increase across all cars that bunch at the notches in table 1 is 109.62 kg.

To estimate the externality associated with increased weight, we use estimates from Anderson and Auffhammer (Forthcoming), which concludes that an increase in vehicle weight of 1000 pounds (454 kg) is associated with a 0.09 percentage point increase in the probability that the vehicle is associated with a fatality, on a mean probability of 0.19%. We use a standard estimate of the value of a statistical life (VSL) of $7 million. Both of these estimates come from the U.S., which is an obvious weakness of the calculation, but we do not know of similar estimates for Japan.

We thus calculate the welfare loss, per car sold, for a 110 kg weight increase as: 110 * 0.0009 * (2.2/1,000) * $7 million = $1525 per car that changes weight in response to the policy. Our data suggest that 9% of cars increase weight and bunch at a notch point. The Japanese car market sells around 5 million new cars per year, so we estimate our aggregate annual welfare distortion to be 9% of 5,000,000 times $1525, which is $686 million.

Our theory showed that there are several distortions caused by attribute-basing. One is the exacerbation of weight-based externalities, which we showed will be "first-order" and thus is expected to dominate the welfare calculation when actual policies are "small" deviations from the first-best. There is also a "Harberger triangle" distortion in the choice of vehicle weight and a distortion due to "general equilibrium" changes in the choice of fuel economy. The existence of these other distortions means that the welfare loss due to accident externalities is a lower bound on the total welfare loss from attribute-basing, because it does not count these other distortions.

In our particular context, the Harberger triangle due to the misallocation of weight may be small. The reason is that consumers presumably do not value vehicle weight directly. Rather, weight is a proxy for other features of a vehicle, as heavier vehicles are larger and have more equipment installed. If an automaker wished to increase the weight of a vehicle holding all else constant, they could presumably do so by adding superfluous materials to the vehicle at relatively low cost (e.g., they could put a lead bar in the trunk). If automakers in fact use this "cheap weight", then the distortionary costs to weight allocation may be as small as the cost of the cheapest available metal, and there would in fact be no Harberger triangle from weight misallocation. Note that "cheap weight" has the same safety implication as "real weight", so that this compliance strategy has no bearing on the weight externality consequences derived above.

Furthermore, in this scenario, we can put an upper bound on the utility lost due to externalities associated with the "general equilibrium" distortion to fuel economy that results from having cars that are heavier than is optimal. In the "cheap weight" scenario, there is no reason to think that consumer utility over fuel economy is affected by the change in weight. There is, however, a direct effect of increased weight on fuel economy, so that the marginal cost curve of providing fuel economy to a vehicle is shifted out when that vehicle becomes heavier. This shift out in the cost schedule will result in a change in private benefits and a change in social benefits due to a change in the amount of externality reduction achieved by the policy. We can bound the latter by assuming that all other features of the car are unaffected, so that the change in fuel economy that results from an increase in weight is due to the mechanical relationship between weight and efficiency.

In particular, the deadweight loss relative to the first-best from reduced fuel economy related externalities per car is equal to the change in fuel economy induced by the change in weight times the change in gallons of gasoline consumed as a result of the change in fuel economy times the externality per gallon. That is, $DWL \approx \varepsilon_{kg}^{mpg} * \frac{\Delta kg}{\bar{kg}} * e_g * g(m\bar{p}g)$, where e_g is the externality per gallon and $g(m\bar{p}g)$ is the mean gallons of gasoline consumed per car.

To calculate this we interpret the estimates in Knittel (2011) as reflecting the engineering relationship. He finds a fuel economy elasticity of -.4 with respect to weight. The mean vehicle weight in our data is around 1400 kg, and our estimated change in weight is 110 kg. At the mean fuel economy in Japan of 14 km/liter (33 mpg), a vehicle that drives 170,000 miles in its lifetime will consumer 5,500 gallons of fuel. We use an aggressive measure of the eternality per gallon of \$2 per gallon, which comes from combining canonical estimates in Harrington, Parry, and Walls (2007) and the accident-related externalities per gallon consumer from Anderson and Auffhammer (Forthcoming). Plugging in these values yields $DWL \approx -.4 * 0.08 * \$2 * 5,500 = \$346$. This is an upper bound in the sense that it assumes no adjustment in technology or other attributes is made in response to the increase in fuel economy and in that it uses an aggressive estimate of the externality per gallon. But, even so, we see that this estimate, while substantial, is less than a quarter of the estimate we get from our weight externality estimate. This suggests that, while the weight externality component that we emphasize is an underestimate of the full welfare costs of attribute-basing, in our case, it likely captures the lion's share of the welfare impacts. This is noteworthy in that we suspect many analysts would intuitively expect the main welfare impacts of attribute-basing to be tied to changes in the externality reduction related to fuel economy.

4 Conclusion

This paper explores the economic implications of attribute-based regulation. We develop a theoretical framework that highlights conditions under which attribute basing is inefficient, and we show that, under those conditions, the use of attribute-based regulation leads to distortions that are concentrated in the provision of the attribute upon which targets are based. The model also explores cases where attribute-basing may play a role, but emphasizes that, even in those cases, the attribute-basing function deviates fundamentally from those observed in real world policies.

Empirically, the paper demonstrates that distortions in response to attribute-based fuel economy standards in Japan are clearly present. We use both established cross-sectional tools based on the notch literature, as well as novel panel techniques, to demonstrate that the Japanese car market has experienced a notable increase in weight in response to attribute-basing.

The theoretical framework makes a number of simplifying assumptions that could be relaxed in future work. In particular, it is possible that attribute-based regulation, particularly when it features notches, impacts firm pricing strategies for certain types of products. Further exploration of the cases in which attribute-basing may be justified is also warranted. On the empirical side, evidence of the responsiveness of attributes other than vehicle weight, which may be particularly easy to manipulate, would be a valuable object of study for future research.

References

Anderson, Michael and Maximilian Auffhammer. Forthcoming. "Pounds that Kill: The External Costs of Vehicle Weight." *Review of Economic Studies* .

Anderson, Soren T., Ian W.H. Parry, James M. Sallee, and Carolyn Fischer. 2011. "Automobile Fuel Economy Standards: Impacts, Efficiency and Alternatives." *Review of Environmental Economics and Policy* 5 (1):89–108.

Blinder, Alan S. and Harvey S. Rosen. 1985. "Notches." *American Economic Review* 75 (4):736–747.

Borenstein, Severin. 2013. "A Microeconomic Framework for Evaluating Energy Efficiency Rebound and Some Implications." Energy Institute at Haas Working Paper 242R.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics* 126 (2):749–804.

Gillingham, Kenneth. 2013. "The Economics of Fuel Economy Standards versus Feebates." Manuscript: Yale University.

Harrington, Winston, Ian Parry, and Margaret Walls. 2007. "Automobile Externalities and Policies." *Journal of Economic Literature* 45 (2):373–399.

Holland, Stephen P., Jonathan E. Hughes, and Christopher R. Knittel. 2009. "Greenhouse Gas Reductions Under Low Carbon Fuel Standards?" *American Economic Journal: Economic Policy* 1 (1):106–146.

Ito, Koichiro. Forthcoming. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review* .

Jacbosen, Mark R. 2013. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior." *American Economic Journal: Applied Economics* 5 (3).

Jacobsen, Mark R. and Arthur A. van Benthem. 2013. "Vehicle Scrappage and Gasoline Policy." NBER Working Paper 19055.

Kleven, Henrik J. and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *The Quarterly Journal of Economics* 128 (2):669–723.

Knittel, Christopher R. 2011. "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector." *American Economic Review* 101 (7):3368–3399.

Metcalf, Gilbert E. 2009. "Tax Policies for Low-Carbon Technologies." *National Tax Journal* 62 (3):519–533.

Saez, E. 2003. "The effect of marginal tax rates on income: a panel study of bracket creep." *Journal of Public Economics* 87 (5-6):1231–1258.

Saez, Emmanuel. 1999. "Do Taxpayers Bunch at Kink Points?" *National Bureau of Economic Research Working Paper Series* No. 7366.

———. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3):180–212.

Saez, Emmanuel, Joel B. Slemrod, and Seth H. Giertz. 2012. "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature* 50 (1):3–50.

Sallee, James M. 2011. "The Taxation of Fuel Economy." *Tax Policy and The Economy* 25:1–37.

Sallee, James M. and Joel Slemrod. 2012. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96 (11-12):981–999.

White, Michelle J. 2004. "The "Arms Race" on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety." *Journal of Law and Economics* 47 (2):333–355.

Whitefoot, Katie and Steven Skerlos. 2012. "Design Incentives to Increase Vehicle Size Created from the U.S. Footprint-based Fuel Economy Standards." *Energy Policy* 41:402–411.

A Our tax model is identical to a regulation

A majority of the attribute-based policies in the real world are regulations, but we wish to study them through the lens of optimal tax policy in order to facilitate comparison to Pigouvian taxes. We demonstrate here the equivalence between regulations and tax policies.

We first specify a fleet-wide energy efficiency regulation in the standard way. Consider a firm that produces goods, indexed by $j = 1, ..., J$ with energy consumption g_j, the sales-weighted average of which is required to exceed a standard σ. Firms can choose each good's price p_j, energy consumption g_j and some other attribute a_j. The firm solves a Bertrand competition problem over all of the products in its portfolio:

$$\max_{g_j,a_j,p_j} \pi = \sum_{j\in J}(p_j - c_j)q_j$$
$$\text{s.t.} \sum_{j\in J}\frac{q_j}{Q}g_j \leq \sigma,$$

where q_j is the quantity sold, which is a function of g, a and p for all vehicles sold in the market; c_j is the marginal cost of vehicle j, which is a function of g_j and a_j but assumed to be constant in q_j, and $Q \equiv \sum_{j\in J} q_j$ is the total sales of the producer across all models. The Lagrangean for this problem is written:

$$\max_{g_j,a_j,p_j} \mathcal{L} = \sum_{j\in J}(p_j - c_j)q_j + \lambda Q\left(\sigma - \sum_{j\in J}\frac{q_j}{Q}g_j\right) = \sum_{j\in J}(p_j - c_j)q_j - \lambda\sum_{j\in J}q_jg_j + \lambda Q\sigma \tag{21}$$

where we multiply the constraint by Q so λ is interpreted as the per-unit shadow cost of tightening the regulatory constraint. This is the standard representation of a fleet efficiency standard.

Alternatively, suppose that there was a tax (or subsidy) equal to $t(g - k)$ per unit sold. That is, there is a marginal tax on fuel consumption g, but we allow also that this marginal tax is shifted uniformly by k so that the net tax on a vehicle could be positive or negative. The profit maximization problem for the firm in this case is:

$$\max_{g_j,a_j,p_j} \pi = \sum_{j\in J}(p_j - c_j)q_j - \sum_{j\in J}q_j(tg_j - k) = \sum_{j\in J}(p_j - c_j)q_j - t\sum_{j\in J}q_jg_j + tQk. \tag{22}$$

If $t = \lambda$ and $k = \sigma$, so that the tax on a product j is $\lambda(g_j - \sigma)$, then the maximization problem for the regulation (equation 21) and the tax problem (equation 22) are identical. Thus, the non-attribute-based regulation is identical to a tax on g.

In an attribute-based regulation, the target efficiency level of each vehicle is a function of the attribute, $\sigma(a)$. The standard that the firm must meet is the average target of its products, which is equivalent to saying that the firm's products must exceed their targets, on average. The Langragean under attribute-based regulation can be written as:

$$\begin{aligned} \max_{g_j, a_j, p_j} \mathcal{L} &= \sum_{j \in J} (p_j - c_j) q_j + \lambda Q \left(- \sum_{j \in J} \frac{q_j}{Q} \left(g_j - \sigma(a) \right) \right) \\ &= \sum_{j \in J} (p_j - c_j) q_j - \lambda \left(\sum_{j \in J} q_j \left(g_j - \sigma(a) \right) \right). \end{aligned} \tag{23}$$

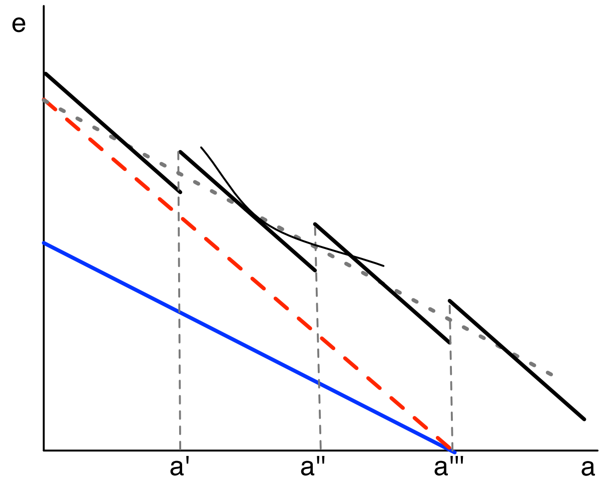
Alternatively, the attribute-based tax policy would be one in which each vehicle faced a tax rate of $t(g - \omega(a))$. In this case, the optimization problem for the firm is:

$$\max_{g_j, a_j, p_j} \pi = \sum_{j \in J} (p_j - c_j) q_j - t \sum_{j \in J} q_j (g_j - \omega(a)). \tag{24}$$

It is immediately apparent then that, when $t = \lambda$ and $\omega(a) = \sigma(a)$, the attributed-based regulation (equation 23) and tax (equation 24) problems are equivalent.

We specified here the problem faced by a single firm. With multiple firms, a common policy function $\sigma(a)$ will correspond to different shadow prices λ, and therefore a different equivalent tax policy. In fully tradable systems (such as CAFE), in which a firm that exceeds the standard can sell its excess credits to another firm, a single shadow price will prevail for all firms, provided that firms do not fail to trade for competitive reasons or because of illiquid markets or transaction costs. Thus, in our baseline model, the modeling of a tax that is uniform across firms embodies the assumption that there is full tradability and no competitive considerations prevent efficient exchange of credits across firms.

Figure 12: Isocost curve with notched attribute-based tax



B Notched policies have similar welfare properties

Many of the attribute-based regulations that we consider feature notches. That is, the function relating the fuel economy target to the attribute, $\sigma(a)$, is a step function with discrete jumps. Sallee and Slemrod (2012) study notched fuel economy policies and conclude that the use of notches in place of a smooth Pigouvian tax is welfare decreasing because it provides overly large incentives for fuel economy improvements for some agents and too small incentives for others. Our context is different in that we have two dimensions, and either could have notches.

We first consider the case where the notches exist only in the attribute function σ, but the tax on e is smooth. That is, conditional on e, the tax policy function $t(a, e)$ takes discrete jumps at certain values of a. How would such a policy function affect consumer choice? We can get some initial intuition graphically. Figure 12 shows an isocost curve, that is, the set of values for which a consumer spends a constant amount on the durable, $P(a, e) + t(a, e)$. The figure is drawn with several notches, at a', a'' and a'''. The solid blue line (drawn to be linear for the sake of illustration) shows the isocost curve before any policy intervention; and the dashed red line shows the modified isocost curve for the same expenditure on the good when there is a Pigouvian subsidy on e that has no attribute slope.

Next, the dashed grey line represents the isocost curve that would exist under a smooth attribute policy. In the diagram, the grey line is drawn parallel to the original blue line, which represents the case when policy makers draw the attribute slope to match existing isocost curves, thereby preserving the original relative prices of a and e. This grey dashed line is not the final isocost curve, however, when $\sigma(a)$ is notched. In that case, the solid black lines represent the isocost curve for the consumer.

Importantly, the line segments on the final isocost curve are parallel to the red dashed line representing the Pigouvian tax. As in the smooth case, the existence of the attribute function does not distort the price of a relative to x, which means that the distortion in the choice of e will be only the indirect change due to a—it will be driven only by the utility and cost interactions of the optimal choice of e and the distorted choice of a. Furthermore, because the line segments are parallel in slope to the original Pigouvian line (and because we assume quasi-linearity) the choice of a will not be distorted if the consumer is choosing an interior point along one of the line segments. All of the distortion is due to cases where a consumer chooses a'. That is, all of the distortion is evident from those who "bunch" at the notch points. This guides our empirical work below.

We now provide algebraic analysis to flesh out the graphical intuition. For notational ease, we focus on the case with only one notch, at a', above which the tax subsidy jumps by amount $\tau > 0$. Then, the tax function can be written as:

$$t(a,e) = \begin{cases} t \cdot e & \text{if } a < a' \\ t \cdot e + \tau & \text{if } a \geq a'. \end{cases} \tag{25}$$

As above, denote by (a^*, e^*) the bundle chosen by a consumer facing a Pigouvian tax of $t \cdot e$. If the consumer's choice under the smooth attribute policy had $a^* > a'$, then the addition of the notch τ is purely an income effect. It has not changed the marginal price of a or e relative to each other or relative to x. Given quasi-linearity, this means that the durable choice of a consumer with $a^* > a'$ is unaffected by the introduction of a notched attribute policy.

When $a^* < a'$, the consumer will face a discrete choice of maintaining their original allocation or switching to a' exactly. They will not choose $a > a'$. To see why, suppose that they chose a value under the notched policy, call it $\tilde{a}$ strictly greater than a'. Then their optimization problem can be written $\mathcal{L} = U(a,e) + [I + G - P(a,e) + t \cdot e + \tau] + \mu[a - a']$, where there is a budget constraint as well as an inequality constraint that $a \geq a'$. If $\tilde{a} > a'$, then the shadow price on the latter constraint,

μ, is zero. In that case, the first-order conditions of the problem will be exactly the same as in the benchmark case with no attribute notch, which by construction featured an optimal choice of $a^* < a'$.

Thus, the consumer with $a^* < a'$ will either choose $\tilde{a} = a^*$ (and not receive τ) or will choose $\tilde{a} = a'$ exactly. This has the empirical implication that all bunching should come "from the left"—changes in a in response to the notched incentives should always be *increases* in a.

If a consumer chooses a', then their choice of e will solve:

$$\max_e = U(a', e) + [I + G - P(a', e) - t \cdot e - \tau], \tag{26}$$

which has the first order condition $\frac{\partial U(a',e)}{\partial e} = \frac{\partial P(a',e)}{\partial e} + t$. This compares to $\frac{\partial U(a^*,e)}{\partial e} = \frac{\partial P(a^*,e)}{\partial e} + t$ for the Pigouvian case, where $a' > a*$. Thus, just as in the smooth case, the distortion in e depends only the relative second-order curvature of P and U, which will be working in offsetting directions. If they offset perfectly, then $\tilde{e} = e^*$, but this would owe to a coincidence regarding the shape of P and U. More generally, the consumer's choice of e could rise or fall relative to the first-best choice.

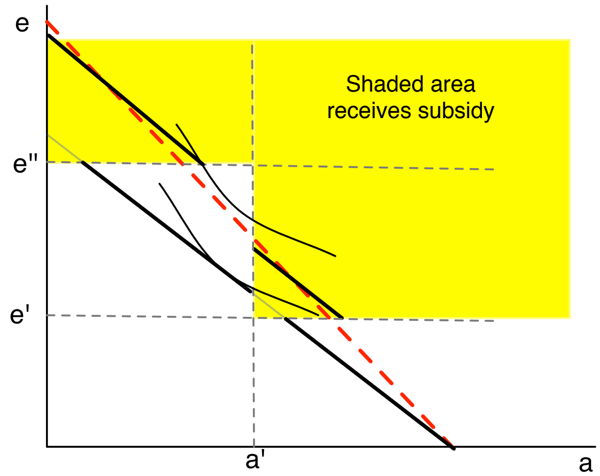
The distortion in a will be analogous to a traditional Harberger triangle and thus rising in τ^2. The consumer will choose $\tilde{a} = a'$ if and only if:

$$-\tau > P(a', \tilde{e}) - P(a^*, e^*) - (U(a', \tilde{e}) - U(a^*, e^*)), \tag{27}$$

that is, whenever the tax benefit is larger than the cost increase from moving from (a^*, e^*) to $(a', \tilde{e})$ minus the increase in utility from that change. The welfare loss can be written as a Taylor expansion, which again has the same intuition as a traditional Harberger triangle.

For our purposes, the main point of this analysis is that, even when the attribute function is notched, the focus of welfare analysis should be on how the policy distorts the choice of a relative to the Pigouvian baseline, and that we should expect the distortion to result in bunching at exactly the notch points in a. For empirical purposes, notched policies are useful in revealing the distortion because it is generally easier to detect bunching at specific notch points than shifts over time in an entire schedule. For that reason, the Japanese fuel economy standards are an especially fruitful context for study.

Figure 13: Isocost curve with notches for both a and e



All of this applies to situations where the (implicit or explicit) tax on e is smooth, but the attribute-basing function is notched. This describes fuel economy regulations in Japan. It is also possible, however, that both the tax on e and the attribute function σ have notches. That is, $t(a,e)$ is discontinuous in both a and e. This case describes the Japanese tax subsidy programs.

The simplest version of this policy is one with a single cutoff for a, call it a' and a pair of cutoffs for e, call them e' and e''. The tax for such a system can be described algebraically as:

$$t(a,e) = \begin{cases} \tau_1 & \text{if } e > e'' \\ \tau_2 & \text{if } e'' > e > e' \text{ and } a > a' \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

An isocost curve for this case is shown in Figure 13. The unsubsidized budget constraint is drawn as a faint line. The final budget constraint is represented by the bold black line segments, which overlap in parts with the unsubsidized line. Allocations in the yellow shaded area receive some subsidy. The subsidy is equal to τ_1 for any allocation above e''. Note that there are large regions of dominance in this diagram, where a subsidized point that has more of a and more of e has the same cost to the consumer as an unsubsidized bundle.

In the diagram, the red dashed line represents the simple Pigouvian tax. The values of τ_1 and τ_2 are chosen in this case to match the average Pigouvian subsidy for the relevant line segments, but this need not be the case. Note that, if it is the case, then $\tau_1 \neq \tau_2$. In many policy examples, $\tau_1 = \tau_2$, which may be suboptimal.

When there are notches in both dimensions, there can be bunching in the distribution of e, at e' and e''. Above we argued that any change in a caused by attribute-basing relative to the Pigouvian optimum would come from *increases* in a. But, in cases with notches in both dimensions, it is possible that responses to the policy will lower a by inducing bunching at e' or e''. This would occur for cases like those represented by the sample utility curve in Figure 13, where a consumer's response to the notched subsidy is to bunch at e''. In that example, the indifference curve that is tangent to the unsubsidized budget constraint features a higher initial choice of a than at the bunch point.

C Technology spillovers create a role for attribute-basing, but not the one given to it by policymakers

Attribute based standards eliminate (or at least limit) the ability of producers to improve energy efficiency by downsizing in a, instead requiring them to make improvements in the products by dialing down attributes other than a that are correlated with e, or by adding technology that improves e while holding a constant. This approach could be appealing if there are positive externalities related to technological advancement. If firms cannot fully appropriate the value of their technological improvements, then technological improvements may represent a spillover benefit to other firms. In this case, there will be too little technological innovation in the absence of policy, and attribute standards might be beneficial in spurring innovation (instead of downsizing).

A direct way to model technological spillovers is to allow some secondary externality to directly enter the welfare function. We develop here a very reduced form model of these spillovers by simply specifying that costs incurred over some reference level, meant to represent the current state of the art or the current production possibility frontier, induce a positive external benefit. We are developing a more complete model of this spillover process, but we strongly suspect that the basic insights of the simplistic model represented here will carry forward in a more rigorous framework.

Denote the frontier cost function (some reference cost level) as $C(\bar{a}, \bar{e}) \equiv \bar{C}$. When the firm chooses a bundle inside of the frontier, that is $C(a, e) \leq C(\bar{a}, \bar{e})$, then the firm creates no technology spillover. The idea in this case is that the good is being produced with already established techniques. If the firm must push costs beyond the current frontier to make a product with a "higher" bundle of a and e, this requires some additional investment that costs $\rho(C(a, e)-C(\bar{a}, \bar{e}))$. We make the functional form assumption that this extra cost is linear and increasing in its argument, so that the extra cost can be written as $\rho(a, e) = \rho \cdot \max\{C(a, e) - C(\bar{a}, \bar{e}), 0\}$ and $\rho > 0$. Conversely, there is some externality that benefits all of society from the technological advancement, which we write as $\gamma(a, e) = \gamma \cdot \max\{C(a, e) - C(\bar{a}, \bar{e}), 0)\}$

The consumer's and planner's problems can be written to include this spillover. There will be first-order conditions for each of two cases, where $C(a, e) > C(\bar{a}, \bar{e})$ and where $C(a, e) \leq C(\bar{a}, \bar{e})$. In either case, the consumer will fail to recognize the externality γ and will recognize only the costs. The only difference in this problem from the baseline model is that there will be a kink in the cost function at $C(\bar{a}, \bar{e})$ which will induce "bunching"—not at a particular level of a or e, but along a particular isocost curve of a and e.

When the first-best choice involves values of a and e below $\bar{C}$, then the consumer's and planner's problems will be identical to the baseline. It is only when the first best allocation lies above $\bar{C}$ that this problem yields different conditions, so we focus on that here. In this case, the planner's problem becomes:

$$\max_{a,e} W = U(a, e) + [I - C(a, e) - \rho \cdot (C(a, e) - C(\bar{a}, \bar{e}))] + \phi \cdot e + \gamma \cdot (C(a, e) - C(\bar{a}, \bar{e})).$$

The planner's first-order conditions of this problem are:

$$\begin{aligned} \frac{\partial W}{\partial a} &= \frac{\partial U}{\partial a} - (1 + \rho - \gamma)\frac{\partial C}{\partial a} = 0 \\ \frac{\partial W}{\partial e} &= \frac{\partial U}{\partial e} - (1 + \rho - \gamma)\frac{\partial C}{\partial a} + \phi = 0. \end{aligned}$$

The consumer's first-order conditions are:

$$\begin{aligned} \frac{\partial W}{\partial a} &= \frac{\partial U}{\partial a} - (1 + \rho)\frac{\partial C}{\partial a} + t\sigma' = 0 \\ \frac{\partial W}{\partial e} &= \frac{\partial U}{\partial e} - (1 + \rho)\frac{\partial C}{\partial e} - t = 0. \end{aligned}$$

To make the consumer's first-order condition for e equal to the planner's, the tax on e must be set at $t = -(\phi + \gamma \frac{\partial C}{\partial e})$. This again is directly consistent with the Pigouvian tax result, where now the total externality related to e includes both the energy externality ϕ and the spillover externality γ. If $t = -(\phi + \gamma \frac{\partial C}{\partial e})$, then to make the consumer's first-order condition for a equal to the planner's, the attribute slope must be:

$$\sigma' = -\frac{\gamma \frac{\partial C}{\partial a}}{\phi + \gamma \frac{\partial C}{\partial e}}. \tag{29}$$

This result indicates that attribute-basing does have a potential role in the first-best solution when their are technology spillovers. Specifically, at the first-best, the slope of the attribute function should be equal to the ratio of the externalities associated with a and the externalities associated with e.

As we argued in section **??**, policymakers appear to commonly choose attribute-based schedules that follow the production frontier. That is, they choose a slope of $\frac{\partial C}{\partial a} / \frac{\partial C}{\partial e}$. This is optimal only if there is no externality related to e, and instead there is only a technological spillover, in which case $\phi = 0$ and equation 29 collapses to be only the ratio of marginal costs of the two attributes. More generally, the larger is the technology spillover γ relative to the energy externality ϕ, the closer will be the attribute slope to the production ratio. In contrast, as the energy externality is greater relative to the technology spillover ($\phi >> \gamma$), the attribute slope will approach zero.

Importantly, this result does not imply that attribute-basing is essential. Instead, attribute-basing provides a second instrument in addition to the tax on e. With two instruments and two externalities (one for e and one for a) it is possible to achieve the first best. That the two instruments are related via attribute-basing is not at all essential. The first best could equally well be achieved by subsidizing e at rate $\phi + \gamma \frac{\partial C}{\partial e}$ and subsidizing a at rate $\gamma \frac{\partial C}{\partial a}$.