

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

Koichiro Ito

James M. Sallee\*

University of Chicago and NBER

University of California, Berkeley and NBER

December 14, 2016

## Abstract

We develop a framework to study “attribute-based regulations,” under which regulatory compliance of a firm, product, or individual depends upon a secondary attribute that is not the intended target of the regulation. Many important economic policies are attribute-based, but economists have given the phenomenon little attention. Our model characterizes the welfare consequences of attribute basing, including its distortionary costs and potential benefits. We empirically test the model predictions by exploiting quasi-experiments in fuel-economy regulations that have fuel-economy targets which are declining step functions of vehicle weight. Using bunching analysis, we identify that the attribute-based fuel-economy regulation induced significant distortionary increases in vehicle weight. Our model also shows conditions under which attribute basing provides efficiency or distributional benefits, at the cost of distorting the attribute. We empirically quantify this welfare trade-off by developing counterfactual simulations based on incentives created by “double-notched” policies.

Keywords: fuel-economy standards, energy efficiency, corrective taxation, notches, bunching analysis

JEL: H23, Q48, Q58, L62

---

\*Ito: [ito@uchicago.edu](mailto:ito@uchicago.edu). Sallee: [sallee@berkeley.edu](mailto:sallee@berkeley.edu). The authors would like to thank Kunihito Sasaki for excellent research assistance. For helpful comments, we thank Hunt Allcott, Soren Anderson, Severin Borenstein, Meghan Busse, Raj Chetty, Lucas Davis, Francesco Decarolis, Meredith Fowlie, Don Fullerton, Michael Greenstone, Mark Jacobsen, Damon Jones, Hiro Kasahara, Kazunari Kainou, Ryan Kellogg, Ben Keys, Christopher Knittel, Ashley Langer, Bruce Meyer, Richard Newell, Matt Notowidigdo, Ian Parry, Mar Reguant, Nancy Rose, Mark Rysman, Jesse Shapiro, Joel Slemrod, Reed Walker, Sarah West, Katie Whitefoot, Florian Zettelmeyer and seminar participants at the ASSA meetings, Berkeley, Boston University, Chicago, the EPA, Harvard, Michigan, Michigan State, MIT, the National Tax Association, the NBER, the RIETI, Pontifical Catholic University of Chile, Stanford, Resources for the Future, UCLA, University of Chile, and Wharton. Ito thanks the Energy Institute at Haas and the Stanford Institute for Economic Policy Research for financial support. Sallee thanks the Stigler Center at the University of Chicago for financial support.  
©2015 Koichiro Ito and James M. Sallee. All rights reserved.

# 1 Introduction

Many important economic policies feature “attribute basing”. An attribute-based regulation aims to change one characteristic or behavior of a product, firm, or individual (the “targeted characteristic”), but it takes some other characteristic or behavior (the “secondary attribute”) into consideration when determining compliance. For example, Corporate Average Fuel Economy (CAFE) standards in the United States are designed to increase the fuel economy of cars (targeted characteristic), but they take the size of each car (secondary attribute) into consideration. As illustrated in Figure 1, firms making larger cars are allowed to have lower average fuel economy. Fuel-economy regulations are attribute-based in the world’s four largest car markets—China, Europe, Japan and the United States. Nearly every wealthy country regulates the energy efficiency of household appliances, and these regulations are universally attribute-based.<sup>1</sup> Consumer-facing product labels are often attribute-based.<sup>2</sup> Regulations ranging from the Clean Air Act to the Family Medical Leave Act, the Affordable Care Act, securities regulations and agricultural regulations are attribute-based because they exempt some firms based on size.

The goal of this paper is to investigate the welfare consequences of attribute-based regulations (ABR), as opposed to regulations based only on the targeted characteristic. Despite the ubiquity of attribute-based policies, the economics literature has not established theoretical and empirical frameworks for the analysis of this important class of policies. In this paper, we first develop a theoretical model that identifies the key parameters that determine the costs and benefits of attribute basing. We then explore two empirical methods that enable us to estimate those key parameters. Our empirical analysis exploits quasi-experiments in attribute-based Japanese fuel-economy regulations, the features of which provide several empirical advantages for estimating the welfare effects of ABR.

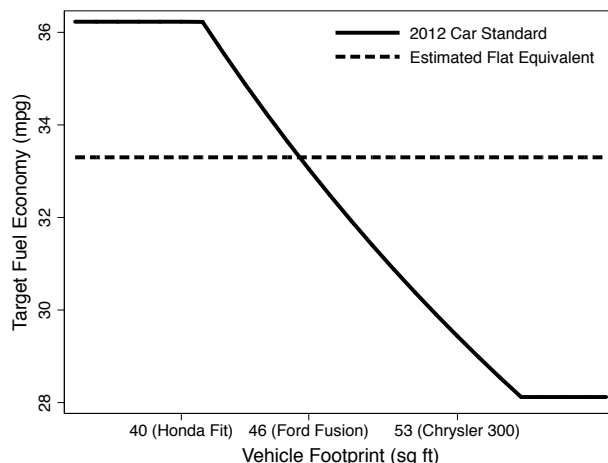
In brief, we conclude that it is unlikely that attribute basing is justified on efficiency grounds. We identify conditions under which attribute basing has efficiency benefits, but these same benefits could be achieved through compliance trading schemes without incurring distortions associated with attribute basing. Attribute basing is an imperfect substitute for compliance trading, and it is justified only if there is some constraint that prevents trading. Instead, we suspect that many ABR are motivated by distributional considerations. In this case, the distortions induced by attribute-basing represent the cost of achieving desired redistribution.

---

<sup>1</sup>For example, a refrigerator in the United States must meet a minimum efficiency that depends on its fresh food capacity and frozen food capacity, as well as its door type (French or not), the location of its freezer (top or bottom), and whether or not it has through-the-door ice.

<sup>2</sup>For example, in Europe, automobiles and appliances are given attribute-based letter grades. In Japan, appliances are given one through five stars, based on an attribute-based criterion. In the United States, energy labels for both automobiles and appliances (including Energy Star certification) include figures that compare the product’s energy consumption relative to products in the same “class”.

**Figure 1:** Example of an Attribute-Based Regulation: U.S. CAFE standard 2012



Solid line depicts the target function, which is a downward-sloping function of vehicle size, as measured by footprint (the square area trapped by the vehicle’s tires). Firms must comply with the target function on average across their fleet. Example vehicles labeled on x-axis.

We begin by establishing a simple theoretical model that facilitates the analysis of attribute-based regulations. The key cost of an ABR is that it creates an implicit incentive for market participants to manipulate the secondary attribute. We argue that this cost, and a variety of possible benefits, can be understood by thinking of the targeted characteristic and the secondary attribute as two distinct goods, the former of which causes an externality. In this framework, the insights and tools of traditional public finance immediately apply. Specifically, our first proposition establishes sufficient conditions under which attribute basing is purely distortionary, because a policy based only on the targeted characteristic can emulate a first-best Pigouvian tax. The resulting welfare loss from attribute basing in this situation is a Harberger triangle in the “market” for the secondary attribute, and thus the elasticity of the attribute with respect to implicit regulatory incentives is the pivotal parameter that determines the magnitude of welfare losses from ABR.

We then investigate a variety of possible benefits of ABR that might rationalize its use despite this cost. We focus on two possibilities, though we discuss several others. First, ABR can enhance efficiency by equalizing marginal costs of regulatory compliance across sources, in certain settings. Some policies (including CAFE) have a compliance trading system, which means that the market as a whole must meet the standard on average, and the market for compliance credits will equalize marginal costs of compliance. Other policies (like energy efficiency mandates for appliances) require each product to comply with a given standard. When each product must comply, marginal costs of compliance will vary across products. When the secondary attribute is correlated with compliance

costs, an ABR can reduce the dispersion in marginal costs. This creates an efficiency benefit that must be weighed against the costs induced by distortions in the choice of the attribute. Our second proposition characterizes this trade-off. Taken together, our results imply that whether an ABR is preferable to a standard based only on the targeted characteristic will depend upon the elasticity of the attribute with respect to regulatory incentives and the degree of marginal cost equalization that the ABR is able to achieve.

A second possibility is distribution. Attribute basing can achieve distributional goals when the planner wishes to shift welfare across consumers or producers based on the secondary attribute. In this case, the efficiency costs of ABR that are our focus represent the cost of achieving distributional goals. For example, size-based fuel-economy regulations can be rationalized as a way of shifting welfare between firms that sell small vehicles and those that sell large vehicles (perhaps in order to favor domestic producers and their consumers). Our final proposition demonstrates conditions under which second-best policies will include attribute basing to achieve redistribution.

In the second part of our paper, we develop two complementary empirical methods that use quasi-experimental policy variation to identify key parameters necessary for assessing the costs and benefits of attribute basing. To do so, we analyze Japanese fuel-economy regulations, under which firms making heavier cars are allowed to have lower fuel economy. The Japanese regulation offers two empirical advantages over data from other markets, including automobile markets in the E.U. and U.S. First, the Japanese regulations have existed for more than three decades and have experienced several policy reforms. Second, the Japanese ABR is notched—the fuel economy required for a given vehicle is a decreasing *step function* of its weight. Automakers therefore have a large incentive to increase vehicle weight only up to key thresholds where the mandated fuel economy drops discretely. These notches do not change the fundamental economic incentives at play, but they aid empirical identification.<sup>3</sup>

Our first empirical strategy is to test for “bunching” (excess mass) in the distribution of vehicle weight around regulatory thresholds, which belie a distortion in vehicle weight (the secondary attribute). We find stark evidence of weight manipulation. Qualitatively, this implies that vehicle weight is responsive to policy incentives, which, according to our theory, implies significant dead-weight loss. To quantify this bunching, we use methods recently developed in the public finance literature.<sup>4</sup> We estimate that 10% of Japanese vehicles have had their weight increased in response to the policy. Among the affected vehicles, we estimate that weight rose by 110 kilograms on aver-

---

<sup>3</sup>In studying notched policies aimed at externalities, our work relates to the literature on notched corrective taxation, which began with Blinder and Rosen (1985), includes prior analysis of automobile fuel economy in Sallee and Slemrod (2012), and is surveyed in Slemrod (2010). Our panel analysis differs from existing work in this area by considering a double notch (i.e., a notch in two coordinates), which is, to the best of our knowledge, new to the literature.

<sup>4</sup>See Saez (1999, 2010); Chetty, Friedman, Olsen, and Pistaferri (2011); Kleven and Waseem (2013); Kleven, Landais, Saez, and Schultz (2014); Decarolis (2015); Fack and Landais (2015) and Gelber, Jones, and Sacks (2015).

age. This not only works against the goal of petroleum conservation (because heavier cars are less fuel-efficient), but it also exacerbates accident-related externalities (because heavier cars are more dangerous to non-occupants). Our back-of-the-envelope estimate based on the value of a statistical life and estimates of the relationship between fatalities and vehicle weight suggests that this weight increase creates around \$1 billion of deadweight loss per year in the Japanese car market. This should spark concern about the use of ABR not only in the substantial automobile market in Japan, which includes roughly five million units sold per year, but also for China (the world’s largest car market), the European Union and India, all of which feature weight-based fuel-economy regulations.

Our second empirical strategy involves estimating a model that enables us to study counterfactual policies and to directly compare the costs and benefits of attribute basing. Our theory emphasizes that ABR can be beneficial in equalizing marginal costs of compliance, in particular for policies that require each individual product to comply with a standard. In such cases, the benefits of marginal cost equalization may outweigh costs from distorting the attribute. Such benefits are likely muted in the Japanese fuel-economy regulations because they allow fleet averaging, but in 2009, the Japanese government introduced a model-specific (rather than corporate average) subsidy for vehicles that exceeded a more aggressive weight-based fuel-economy threshold. This provides an ideal opportunity to use quasi-experimental variation and revealed preference data to estimate parameters necessary for directly comparing the efficiency benefits and distortionary costs of an ABR. Vehicles that are modified in order to become eligible for the subsidy reveal information about the relative costs of changing weight versus fuel economy. We construct panel data spanning the introduction of the subsidy, and use it to estimate the compliance costs of modifying fuel economy and weight.

We use these estimates to evaluate three counterfactual policy scenarios—attribute-based fuel-economy standards, a flat standard without compliance trading, and a flat standard with compliance trading. Consistent with the results of our model, when compliance trading is disallowed, attribute-based standards improve efficiency as compared to a flat standard because attribute basing helps equalize marginal compliance costs. However, this benefit is partially offset by distortions in the attributes created by the regulatory incentive; the ABR results in weight increases, whereas a flat policy leads to weight reductions. Also consistent with our theory, we find that attribute basing is an imperfect substitute for compliance trading because marginal compliance costs are not perfectly correlated with the attribute, which results in only partial equalization of the marginal compliance costs. In our case, the ABR recovers only about half of the welfare gain that would be achieved by a flat standard with compliance trading.

Our paper contributes to several literatures. First, we contribute to the environmental economics literature on differentiated standards, of which attribute basing is an example. The prior literature has

focused on vintage-differentiated standards (e.g., Gruenspecht 1982; Nelson, Tietenberg, and Donihue 1993; Stavins 2006) or spatial differentiation (e.g., Mendelsohn 1986; Becker and Henderson 2000). Consistent with our findings, the literature on spatial differentiation finds that such differentiation may be logical from a cost-effective standpoint, but that it may distort the location of economic activity. Similarly, the literature on vintage differentiation notes that differentiation can improve efficiency by equalizing marginal costs of compliance, but that this comes at a cost of distorting the vintage distribution by affecting exit or entry. Parallel to our conclusions, the literature suggests that tradable permits may be a superior way to achieve these gains (see, for example, Stavins 2006). Relative to this literature, our model offers a concise way of characterizing benefits and distortions, as well as optimal policies, in a unified welfare framework. We also extend the literature to consider differentiation in product characteristics like size.

Second, we contribute to the analysis of fuel-economy policies. There is a substantial literature in this area, but few studies have considered attribute basing. Whitefoot and Skerlos (2012) use engineering estimates of design costs and a discrete-choice model to predict the manipulation of footprint in response to CAFE. Reynaert (2015) studies the roll out of fuel-economy standards in the E.U., which are weight-based, using a structural model of the market. Relative to those papers, we provide revealed preference evidence of weight changes using quasi-experimental variation. Our paper is also the first to develop a full theoretical model of attribute basing, though Gillingham (2013) notes the potential implicit incentive for the expansion of footprint in a broader discussion of CAFE policies. Finally, Jacobsen (2013b) addresses the safety impacts of footprint-based standards in the United State, which we discuss below.

Third, there is a literature on regulations that exempt firms based on size. Empirical results from this literature have been mixed (Becker and Henderson 2000; Gao, Wu, and Zimmerman 2009; Sneeringer and Key 2011). The theoretical literature has focused on models of imperfect targeting of the externality (Brock and Evans 1985; Kaplow 2013) or per-firm administrative costs (Dharmapala, Slemrod, and Wilson 2011). In contrast, we find stark empirical evidence of bunching, and we consider efficiency benefits due to potential marginal cost equalization, as well as redistribution.

In sum, our paper demonstrates how ABR, which may at first appear quite different from a tax, can be analyzed with the tools of public economics—i.e., Pigouvian taxes, Harberger triangles, and the theory of the second best—to evaluate an important class of policies. For the policy we study, we find large unintended consequences, and we quantify the benefits of alternative policies. Our paper has important implications for future research and environmental policy because a growing number of countries are adopting attribute-based regulations in durable goods markets, including automobiles, electric appliances, solar panels and buildings.

## 2 Theory

Our model setup is as follows. A consumer has unit demand for a durable good with two continuously varying characteristics  $a$  and  $e$ , the levels of which they choose. The present discounted benefit of services from the durable is  $F_n(a_n, e_n)$ , where  $n = 1, \dots, N$  indexes a type of consumer whose tastes may vary. Consumers have exogenous income  $I_n$ , which they spend on the durable and a quasi-linear numeraire  $x$ . The characteristic  $e$  generates an externality that is linear in the aggregate  $e$  consumed over all types, with marginal social benefit  $\phi$ . The total externality is  $\phi \sum_n e_n$ . In our terminology,  $e$  is the targeted characteristic;  $a$  is the secondary attribute.

We model an attribute-based regulation as a mandate that requires  $e_n \geq \sigma(a_n)$ . This mandate acts as a constraint on the consumer’s optimization problem. When compliance trading is allowed, the mandate must be met by the market on average, but individual products can make up a compliance gap by purchasing credits.<sup>5</sup> We generally work with a linear attribute-based regulation, which has  $e_n \geq \hat{\sigma}a_n + \kappa$ , where  $\hat{\sigma}$  and  $\kappa$  are constants. Where we work with a more general function  $\sigma(a)$ , we assume it is differentiable and includes a constant term  $\kappa$ . We call  $\sigma'(a_n)$ , which equals  $\hat{\sigma}$  for linear policies, the “attribute slope”.<sup>6</sup>

We assume a perfectly competitive supply side with no fixed costs per variety. This means that consumers can choose any bundle of  $a$  and  $e$  and pay a price  $P(a, e)$  that is equal to the marginal cost of production  $C(a, e)$ , which we assume is increasing and convex in both arguments. Our supply assumptions simplify the analysis and allow us to focus on the unique implications of attribute basing. There are no firm profits, and there is no distinction between changing the attributes of an existing product versus introducing a new variety. In our empirical application, we observe products made by firms, and to apply our model we interpret changes in product characteristics as reflecting policy-induced shifts in consumer choices. The benefit is tractability and transparency, but the cost is that we do not account for the welfare consequences of policy-induced changes in firm market power.

Under marginal cost pricing, having substituted in the budget constraint, consumer  $n$ ’s Lagrangean is:

$$\max_{a_n, e_n} U_n = F_n(a_n, e_n) - C(a_n, e_n) + I_n + \lambda_n(e_n - \sigma(a_n)), \quad (1)$$

---

<sup>5</sup>Some attribute-based regulations feature fleet averaging, where all of the products sold by a particular firm must comply on average. In terms of welfare impacts, fleet averaging is an “in between” case, which we discuss further in section 2.4.

<sup>6</sup>The constant  $\kappa$  implies that a linear policy can correct average prices, which distinguishes our policy from some performance standards. We discuss this issue further below. In second-best cases, a nonlinear tax may offer some efficiency gains, but the linear assumption helps simplify the analysis. Our theory focuses on differentiable (smooth) policies, but our empirical analysis considers a case where  $\sigma(a_n)$  is a step function. The welfare implications of attribute basing are quite similar for such policies, which we discuss in an appendix, and we present the smooth policy version believing it to be more intuitive and general.

where  $\lambda_n$  is the shadow price of the regulation. The consumer ignores the externality when making choices.

The planner puts welfare weight  $\theta_n$  on the utility of type  $n$ , which includes the externality, where the mean of  $\theta_n$  is normalized to 1. The planner maximizes social welfare  $W$  by choosing the policy function  $\sigma(a_n)$ :

$$\max_{\sigma(a_n)} W = \sum_n \theta_n (F_n(a_n, e_n) - C(a_n, e_n) + I_n) + \phi \sum_n e_n. \quad (2)$$

At times, we make use of notation that writes the consumer's welfare loss from deviating from their private optimum as  $L_n(e_n - e_n^0, a_n - a_n^0) \equiv U_n(a_n, e_n) - U_n(a_n^0, e_n^0)$ , where  $a_n^0$  and  $e_n^0$  are the characteristics that type  $n$  would choose in the absence of any policy (i.e., their privately optimal bundle). We denote deviations from these private optima as  $\Delta a_n \equiv a_n - a_n^0$  and  $\Delta e_n \equiv e_n - e_n^0$ .

In some cases, we will assume a quadratic functional form of  $L_n$  for illustration:

$$L_n(\Delta a_n, \Delta e_n) = \alpha(\Delta a_n)^2 + \beta(\Delta e_n)^2 + \gamma\Delta a_n\Delta e_n. \quad (3)$$

Our model makes several assumptions in the interest of simplicity, including perfect competition, perfect targeting of the externality, and unit demand for the durable. We discuss these issues further in section 2.4.

## 2.1 Attribute basing is an imperfect substitute for compliance trading

Some attribute-based regulations (including CAFE since 2012) have a compliance trading system through which firms that exceed the standard are given a credit for excess compliance that can be sold to another firm. Buyers can use credits to achieve compliance. If the market for permits is competitive, a trading system ensures that the marginal compliance cost, which will equal the equilibrium price of a compliance credit, is uniform across all firms and products.<sup>7</sup> We first consider the implications of ABR in the presence of such a competitive compliance trading system.

When there is compliance trading, the potential benefit that ABR provides by equalizing marginal costs of compliance is obviated. Proposition 1 shows that optimal policy involves no attribute basing in this case.

**Proposition 1.** *Assume that there is competitive compliance trading. If welfare weights are uniform*

---

<sup>7</sup>One might be concerned about efficiency in such a market with a small number of firms, but Leard and McConell (2015) discuss evidence that trades have taken place in the early years of CAFE trading.



$(\theta_n = 1 \forall n)$ , the optimal policy involves no attribute basing. The optimal attribute slope is:

$$\sigma'(a_n)^* = 0 \forall a_n.$$

The proof of Proposition 1 shows that the first-best allocation is achieved by a flat standard that is set at a level that implies that the market shadow price is equal to  $\phi$ , the externality. (All proofs are in the appendix.) This is intuitive. The crux of our argument is that the bundling of  $a$  and  $e$  in a single durable good is largely irrelevant; the consumer’s problem can be understood as a microeconomic choice problem over two related goods,  $a$  and  $e$ . The only difference is that the price of  $a$  and  $e$  can be nonlinearly related. As can be seen by the first-order conditions of the consumer’s problem (equation 1), attribute-based regulation creates a pair of wedges, equal to  $\lambda$  and  $-\lambda\sigma'(a_n)$ , in the “markets” for  $e$  and  $a$ , respectively.

The conditions of Proposition 1 imply that the planner has no distributional concerns, marginal costs are equalized across types by compliance trading, and the only market failure is the externality from  $e$ . As a result, the planner can achieve the first-best allocation by creating a wedge in the choice of  $e$  equal to the externality. There is no benefit to creating a wedge in the choice of  $a$ , which requires that  $\sigma'(a_n) = 0$ . Where there are other market failures, this result will change, which we discuss in section 2.4.

This is consistent with standard principles of Pigouvian taxation, which we emphasize by restating the result for a tax policy instead of a regulation. Suppose that instead of a regulation with compliance trading there is a subsidy for the durable equal to  $s \times (e_n - \sigma(a_n))$ , which is made revenue neutral through a lump-sum uniform tax collected from all types. Corollary 1 restates Proposition 1 for this case.

**Corollary 1.** *Assume welfare weights are uniform ( $\theta_n = 1 \forall n$ ). The optimal subsidy involves no attribute basing. The optimal attribute slope is:*

$$\sigma'(a_n)^* = 0 \forall a_n.$$

The literature on Pigouvian taxes has long contemplated an “additivity property”, which states that (a) the optimal tax on a commodity that produces an externality is equal to the optimal tax on that good if there were no externality, plus marginal external damages; and (b) the externality does not change the optimal tax on other goods, even if they are substitutes or complements to the externality-generating good. This property holds broadly (Kopczuk 2003). Our result is a manifestation of the additivity property, and will thus hold in a wide range of second-best settings.

When there is not a compliance trading system, each product must individually comply with the mandate. For a flat standard, this will give rise to dispersion in the marginal costs of compliance across products, which violates the equimarginal principle and belies inefficiency. When the attribute is correlated with compliance costs, ABR can have efficiency benefits by reducing this variation. This is shown in Proposition 2.

**Proposition 2.** *Assume that there is no compliance trading. Then, even if welfare weights are uniform ( $\theta_n = 1 \forall n$ ), the optimal linear regulation generally involves attribute basing. If the constraint binds for all  $n$ , the optimal attribute slope satisfies:*

$$\hat{\sigma}^* = \frac{\text{cov}(\lambda_n, a_n)}{\phi \left( \frac{\partial \bar{a}}{\partial \hat{\sigma}} - \bar{a} \frac{\partial \bar{a}}{\partial \kappa} \right)},$$

*which is not zero unless  $\lambda_n$  is uncorrelated with  $a_n$ .*

Proposition 2 shows that, absent a compliance trading system, some attribute basing is optimal, even when there are no distributional considerations. The exception is when the attribute is perfectly uncorrelated with marginal compliance costs under a flat standard, in which case the numerator is zero, and attribute basing, which cannot equalize compliance costs, is undesirable. An ABR is a differentiated standard; so long as the dimension along with the standard can be differentiated is correlated with marginal cost, there is a potential efficiency gain.

Attribute basing is a substitute for compliance trading, but it is an imperfect substitute for two reasons. First, whereas compliance trading can generate first-best outcomes in our framework, attribute basing can improve marginal cost equalization, but only partially, unless compliance costs are perfectly predicted by the choice of  $a$ . Second, attribute basing achieves marginal cost equalization by inducing distortions in the choice of the attribute for all types, which has an efficiency cost. In our empirical analysis below, we estimate a model that allows us to directly compare the distortionary costs and the marginal cost equalization benefits.

We focus our attention on the possibility of using ABR to equalize marginal compliance because this motivation is consistent with the design of several real policies. To illustrate this, we first make a further point about marginal cost equalization in corollary 2 and then discuss how this relates to real world examples. Corollary 2 uses an example to illustrate that, even when an ABR could achieve full equalization of marginal costs, it will not be optimal to do so.

**Corollary 2.** *Assume that there is no compliance trading, that welfare weights are uniform ( $\theta_n = 1 \forall n$ ), that the constraint binds for all  $n$ , and that there is a perfect correlation between attributes ( $e_n^0 = b + ma_n^0$  with  $m \neq 0$ ). With a uniform quadratic loss function for all  $n$ , the optimal linear regulation involves attribute basing but it does not fully equalize marginal costs, even though this is possible. The optimal attribute slope satisfies  $\hat{\sigma}^* \neq 0$  and  $\hat{\sigma}^* \neq m$ .*

The corollary adds two assumptions to Proposition 2—a uniform quadratic loss function and a perfect correlation between the privately optimal bundles,  $e_n^0$  and  $a_n^0$ —that together imply perfect marginal cost equalization is possible.

To see this, note that the quadratic loss function implies that the shadow price of a policy for type  $n$  can be written in closed form as:

$$\lambda_n = \xi \times (\hat{\sigma}a_n^0 + \kappa - e_n^0) \tag{4}$$

$$\text{where } \xi = \frac{4\alpha\beta - \gamma^2}{\beta\hat{\sigma}^2 + \gamma\hat{\sigma} + \alpha}.$$

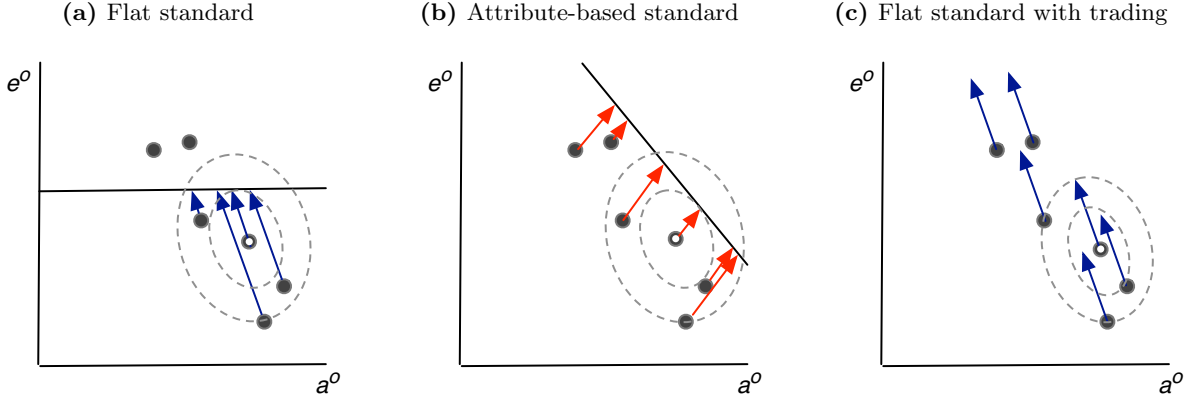
Marginal cost is a function of parameters of the adjustment cost function (equation 3) and the compliance gap at the private optimum. When there is perfect correlation ( $e_n^0 = a + ma_n^0$ ) and  $\hat{\sigma} = m$ , marginal cost is equal to  $\xi \times (\kappa - b)$  for all  $n$ .

The point of corollary 2 is that, even when this is possible, it is not optimal. Beginning from a flat standard, a steeper attribute slope will increase marginal cost harmonization, but it also exacerbates distortions in  $a$  (for all types). The second-best attribute slope strikes a balance between the costs of distorting  $a$  and the benefits of marginal cost harmonization. An ABR is a differentiated standard, and this balancing of costs and benefits will be true in other types of differentiated standards, so long as (a) marginal costs are correlated with the variable that determines differentiation, and (b) economic agents can strategically alter the variable that determines differentiation.

Whether or not there is a perfect correlation between  $a_n^0$  and  $e_n^0$ , for the quadratic case, the planner can minimize the variation in marginal cost by choosing  $\hat{\sigma}$  to minimize  $\sum_n (\hat{\sigma}a_n^0 + \kappa - e_n^0)^2$ . This is directly analogous to minimizing the sum of squared residuals, and the solution will be the Ordinary Least Squares solution. That is, the planner can maximize marginal cost harmonization by choosing  $\hat{\sigma}$  as the best fit slope relating  $e_n^0$  and  $a_n^0$ . Corollary 2 shows that, even when perfect harmonization is possible, it is not optimal; instead, optimal policy should leave more marginal cost variation than is necessary in order to mitigate the distortion to the choice of  $a$ .

This is relevant because real-world attribute slopes have often been determined by fitting actual data, without attention to the second-best considerations we emphasize here. For example, U.S. regulators chose the slope of the footprint-based standard in CAFE by fitting a line to data on fuel economy and footprint. Fuel-economy standards in the E.U. were similarly designed by estimating the relationship between fuel economy and weight. Japanese fuel-economy standards do something similar, though the slope is chosen to fit only a subset of vehicles deemed to be high performing. Thus, policy makers seem to have designed ABR to maximize marginal compliance equalization, but this is not optimal because it ignores the distortion to the attribute. The second-best policy should be a

**Figure 2:** Graphical illustration of ABR



Points represent privately optimal choice for a type. Ellipses (drawn for one point only) are level sets of loss function. Optimal compliance path, which depends on policy, drawn as vector. With uniform quadratic loss function, vectors are parallel across points of a given policy, and vectors are parallel in (a) and (c).

compromise between marginal cost equalization and mitigation of the distortion to the attribute.

### Graphical illustration

We provide a brief graphical illustration of these conclusions in Figure 2, where each dot represents the privately optimal bundle for a type,  $(a_n^0, e_n^0)$ . The three panels of the figure represent three different policies. Vectors depict privately optimal compliance choices; that is, the end point of the vector is the new attribute bundle chosen by that type in order to comply at the lowest possible cost.

By definition, any movement away from the private optimum causes a private welfare loss. For quadratic losses, the level sets of  $L$  will be ellipses surrounding the private optimum. Figure 2 depicts level sets for one data point. In the absence of compliance trading, when faced with a binding regulation, the cost-minimizing way to achieve compliance for each type will be to relocate to a point where the lowest possible level set of the loss function is tangent to the regulation.

Given the quadratic loss function, the length of a compliance vector will be directly proportional to marginal cost. Marginal cost equalization is thus signified when vectors are all the same length. The slope of the privately optimal compliance vector is determined by the attribute-slope. Specifically, compliance vectors will have slope  $(4\alpha\beta - \gamma^2)/(2\beta\hat{\sigma}^2 + \gamma)^2$ , which is a function of  $\hat{\sigma}$ . When  $\hat{\sigma}$  is zero, the compliance vector has slope  $(4\alpha\beta - \gamma^2)/\gamma^2$ , which may be negative or positive; that is,  $a$  might rise or fall in response to a flat standard, depending on the curvature of the loss function.

The three panels of Figure 2 depict the response to a flat standard, an attribute-based standard, and a flat standard with compliance trading. In Figure 2a, a flat standard generates no response (no

vector depicted) among some types, because their private optimum is above the standard. They have a zero marginal cost; the standard is not binding. For products with non-zero marginal costs, the marginal cost varies, which is indicated by the differences in vector lengths.

Figure 2b depicts an attribute-based regulation. It partially equalizes marginal costs (equalizes vector length), more so as the correlation between  $a_n^0$  and  $e_n^0$  is tighter. If the correlation between  $a_n^0$  and  $e_n^0$  was perfect, the equalization could be perfect.

But, attribute basing also induces a change in the slope of the compliance vector. When  $\hat{\sigma}$  is larger (in absolute value), the slope of the compliance vector will become flatter. That is, the proportion of the response to the ABR that comes from changes in  $a$  rather than changes in  $e$  will rise. This change in slope is inefficient, which can be seen by comparison with Figure 2c, which depicts a fully efficient flat standard with compliance trading. The slopes of the vectors are the same in Figures 2a and 2c; both are flat policies and induce a slope of  $(4\alpha\beta - \gamma^2)/\gamma^2$ . But, the compliance trading system equalizes shadow prices so that all products have the same vector length.

Thus, the limitations of attribute basing are twofold. First, an ABR will only partially equalize marginal costs. Second, the ABR achieves marginal cost equalization by distorting the choice of  $a$ . As a result, as indicated by Corollary 2, the second-best attribute-slope will not be the one that maximizes marginal cost equalization. Instead, it will be less steep; it will trade-off mitigation in the distortion to  $a$  for more dispersion in marginal cost. Of course, in reality, the loss function may differ across products, which creates another advantage for the efficient policy relative to a flat standard or an ABR. We explore this possibility in our simulations in section 4.

The graphs also help illustrate a common misunderstanding in the non-academic literature, which implies that it is desirable for an attribute-based standard to not distort the distribution of the secondary attribute *relative to the no-policy baseline*.<sup>8</sup> An efficient policy will induce a change in the attribute (the compliance vectors in Figure 2c are not perfectly vertical), and an attribute-based policy that preserves the distribution of the attribute from a no-policy baseline is preserving a market inefficiency that results from private agents ignoring the externality.

We do not formally analyze the optimal choice of *which* attribute to use in this paper. But it is clear that attributes that are (a) more closely related to compliance costs and (b) less elastic will be better. The former characteristic maximizes the ability of an ABR to harmonize compliance costs; and the latter characteristic minimizes the welfare cost of distortions to the attribute. In the quadratic case, the variance of the shadow price is  $var(\hat{\sigma}a_n^0 - e_n^0) \times \xi^2$ . This means that the  $R^2$  from a regression of  $e_n^0$  on  $a_n^0$  is proportional to the fraction of this variation that can be reduced by a linear ABR. The challenge for policymakers in finding the right attribute is twofold. First, knowledge of marginal costs

---

<sup>8</sup>For example, see the Federal Register, volume 77, number 199, page 62687 for a discussion related to footprint-based CAFE.

will be required to estimate correlations. Second, the correlation between the attribute and marginal cost may be endogenous to the policy regime. Further exploration is a promising topic for future research.

Some real world policies, including EU fuel-consumption standards, CAFE prior to 2012, and the Japanese fuel-consumption regulations that we analyze in section 4, represent an intermediate case by allowing fleet averaging. A firm must comply on average across its products, but there is no trading allowed between firms. Fleet-averaging will harmonize the marginal costs within a firm, but not across firms. Fleet averaging may achieve a substantial fraction of welfare gains. It can be thought of as a constrained form of trading in our model.

In sum, for policies that require each individual product to meet a standard rather than for the market as a whole, an ABR can increase efficiency. Minimum efficiency standards for appliances are an example of this type of policy. The first two phases of Chinese fuel-economy standards were also designed this way, as is the tax subsidy for automobiles in Japan that we analyze in section 3. ABR can increase efficiency for such policies by harmonizing marginal costs, but it will do so imperfectly unless the attribute is perfectly correlated with marginal costs, and it can only do so by creating a distortion in the choice of the attribute, unless that attribute is completely unresponsive to policy incentives. To the degree that the conditions are not met, ABR will be less efficient than a flat standard with compliance trading, but it may still improve over a flat standard without compliance trading. In our empirical analysis, we estimate a quadratic loss function and use it to calculate the relative benefits of an ABR over the flat standard and the relative inefficiency when compared to a flat standard that includes compliance trading.

## 2.2 The deadweight loss of ABR

If attribute basing is employed in a policy with compliance trading (perhaps due to political constraints or simply by mistake), it will create a welfare distortion without any resulting benefit. Specifically, Proposition 3 shows that, for a subsidy, when the subsidy on  $e$  is set to the Pigouvian benchmark, deadweight loss is approximated by the Harberger triangle in the “market” for  $a$ , summed across types.

**Proposition 3.** *Assume welfare weights are uniform ( $\theta_n = 1 \forall n$ ) and  $s = \phi$ . The deadweight loss from a subsidy with  $\sigma'(a_n) \neq 0$  is approximated as:*

$$DWL \approx \sum_n 1/2 \cdot \frac{\partial a_n}{\partial (s\sigma'(a_n))} (s\sigma'(a_n))^2.$$

The wedge in the choice of  $a$  is  $s\sigma'(a_n)$ , so this is directly analogous to a Harberger triangle. Just as

in a standard setting, the magnitude of the welfare loss from using attribute basing is determined by the size of the wedge ( $s\sigma'(a_n)$ ) and the derivative of the good ( $a$ ) with respect to that wedge. As  $a$  is more elastic, distortions from ABR will be larger. The central object of our empirical analysis is to determine whether, for the ABR we study, this derivative is large or small, as that will determine the size of any distortion.

In the appendix, we show also that, if a policy is constrained to include attribute basing, this leads to an attenuation of the second-best tax away from the Pigouvian benchmark of marginal damages, because the distortionary wedge imposed on  $a$  is mechanically related to the wedge in  $e$ .

When the attribute  $a$  is fixed and does not respond to policy, there will be no cost from a distortion in  $a$ . Even in this case, compliance trading offers an information advantage over attribute basing: the design of efficient attribute-based policies requires information about marginal costs and their correlation with the attribute.

### 2.3 Distributional justifications for ABR

The preceding analysis assumes away distributional considerations so that the planner cares only about efficiency. A change in the attribute-slope will affect a transfer of welfare across types according to their demand for the attribute. Thus, when welfare weights are correlated with demand for the attribute across types, distributional considerations will give rise to attribute basing. Proposition 4 demonstrates this for the linear subsidy.

**Proposition 4.** *Assume that welfare weights  $\theta_n$  vary. Then, the optimal linear subsidy involves attribute basing unless  $\theta_n$  is uncorrelated with  $a_n$ . The optimal attribute slope is:*

$$\hat{\sigma}^* = \left[ \left( \frac{\phi - s}{s} \right) \frac{\partial \bar{e}}{\partial \hat{\sigma}} - \text{cov}(\theta_n, a_n) \right] \bigg/ \frac{\partial \bar{a}}{\partial \hat{\sigma}}.$$

The optimal attribute slope has two terms. The first is zero when the subsidy is set equal to marginal damages ( $s = \phi$ ). In that case, the optimal slope is the ratio of the covariance between the attribute and the welfare weight and the average derivative of  $a_n$  with respect to the attribute slope. When the correlation between  $\theta$  and  $a$  is positive, the optimal attribute slope is negative (because  $\partial a / \partial \hat{\sigma}$  is negative). As that correlation gets stronger, the slope becomes steeper. However, as the attribute is more elastic with respect to the policy wedge ( $\partial a / \partial \hat{\sigma}$  is larger in magnitude), the optimal slope is flatter. This is because any distributional gains are achieved at the cost of distorting the choice of  $a$  for all types. When the attribute is more elastic, the efficiency costs of distribution are higher. Less redistribution is therefore optimal, and the optimal slope is flatter. When  $s \neq \phi$ , the first term will be

non-zero. This term shrinks as the elasticity of the attribute, relative to the elasticity of the targeted characteristic, rises.

We suspect that distributional concerns are a key explanation for the use of attribute basing in real policies. For example, it is widely held that footprint-based CAFE standards were designed in part to shift welfare towards the Detroit Three, who make larger cars than their main competitors. Similarly, French and Italian automakers (who make relatively light cars) argued in favor of a flat standard, while German automakers (whose cars are relatively heavy) argued for a weight-based standard in the E.U. Any distributional gains come at an efficiency cost of distorting footprint or weight, more so as the attribute is more responsive to policy.

## 2.4 Additional implications of attribute basing

Our model is focused on what we believe to be the core economic implications of attribute basing: it creates an incentive to distort the secondary attribute, which might be justified by distributional considerations or marginal cost equalization. We modeled those two justifications because we think they are the most likely explanations for actual policies. Under certain circumstances, there could be other benefits to attribute basing. We briefly describe them here but leave formal modeling for future research.

First, our model assumes perfect competition. Imperfect competition implies that the private market is not welfare maximizing, even when the externality is corrected. It is thus conceivable that attribute-based regulation could be used to mitigate distortions due to market power, but we are aware of no evidence that policy makers have ever actually considered this.<sup>9</sup>

Second, we assume unit demand for the durable. A general problem with performance standards is that they may fail to induce the right shrinkage (or expansion) in the overall market. For example, because all cars emit carbon, an efficient carbon tax would shrink the aggregate car market, as all cars would face a positive tax. In contrast, a flat fuel-economy regulation will (implicitly) tax some products while subsidizing others. A flat regulation has only one choice parameter, and it will generally be unable to correct both relative prices (across cars) and average prices simultaneously.<sup>10</sup> An attribute-based regulation can potentially ensure that the price of all products rises, thereby correcting average prices, but an ABR does this by inducing distortions in the choice of the attribute. The same benefits could be achieved without introducing a distortion in the attribute by combining a flat standard with a sales

---

<sup>9</sup>In general, whether an ABR based on other motivations will mitigate or exacerbate market power distortions will depend on the nature of competition and the distribution of preferences. Specifically, based on the logic of Spence (1975), we conjecture that the key issue is whether the marginal consumer for each product values the attribute more or less than average.

<sup>10</sup>Holland, Hughes, and Knittel (2009) treat this issue in detail for the case of a low-carbon fuel standard. Jacobsen and van Benthem (2013) suggest that, for the case of automobiles, the benefits resulting from higher new car prices will be partly offset by changes in vehicle scrappage.



tax or registration fee that shifts all prices equally.

Third, advocates of attribute basing in car markets have argued that it promotes technology adoption.<sup>11</sup> Roughly speaking, automakers can comply with a flat standard by downsizing their fleet or by adopting new technologies. Attribute-based policies can be designed to limit opportunities for downsizing, which forces compliance to come from technology. If there are spillovers between companies from adopting new technologies, there might be some justification for attribute basing. We are, however, skeptical that technological spillovers are large in the auto market, as there is extensive patenting and licensing.

Fourth, attribute-based regulation may actually have some efficiency advantages when uncertainty is introduced. For example, to a first-order approximation, the optimal gasoline tax does not change when the market price of petroleum moves up or down. But, under a flat standard, gasoline price fluctuations will move the shadow price of fuel economy regulation as consumers shift demand between smaller and larger vehicles. The shadow price of a footprint (or weight) based standard will fluctuate less, because it does not depend (or depends less) on the market demand for small versus large vehicles.<sup>12</sup>

Fifth, our model assumes that the externality is produced by  $e$ , which can be targeted directly by policy. In reality, energy-efficiency ratings do not directly cause externalities, and energy-efficiency policies are always therefore second-best instruments, which have well known limitations compared to Pigouvian taxes. Most often discussed in the literature is the fact that energy efficiency policies fail to provide incentives on the intensive use margin (e.g., fuel economy standards induce additional miles traveled, instead of reducing them as would a gasoline tax, by lowering the cost of driving per mile). This is an important limitation of energy efficiency policies, but we believe it is largely orthogonal to attribute basing, which provides no direct way of influencing the utilization margin. An extended explanation of this conclusion is provided in the appendix section.

A related second problem is that energy efficiency is necessarily a noisy proxy for the externality. To see the implications of this for attribute basing, suppose that the externality is a function of  $e$  and some other factor  $\omega$  upon which policy cannot be based. Then, it might be useful to base policy on a secondary attribute  $a$ . Akerlof (1978) coined the term “tags” in a related model of income taxation; the planner wishes to tax ability, but makes taxes contingent on other observable characteristics because

---

<sup>11</sup>Advocates have also argued that attribute basing promotes safety by promoting larger cars. This appears to be based on a misunderstanding of safety-related externalities. Larger cars are safer for the car’s occupants (which is a private benefit and should be priced into the car), but they are more dangerous to those outside the car (which is an externality). If ABR changes the distribution of sizes of cars, this could affect net safety. See Jacobsen (2013b) for a related model that concludes that footprint-based CAFE is roughly safety neutral.

<sup>12</sup>We are especially grateful to Ryan Kellogg for suggesting this line of reasoning. Consistent with this, Leard, Linn, and McConell (2016) suggest that fleet fuel economy appears to be less responsive to gasoline price fluctuations since the introduction of attribute-based standards.

they are correlated with income. The secondary attribute can be a tag in this setting. In this case, the welfare improvement from tagging will depend on the correlation between  $\phi(e, \omega)$  and  $a$ , *conditional on  $e$* . But, actual policies seem to have selected attributes that are tightly correlated with  $e$ , which limits the usefulness of the tag. For example, when discussing the decision to use footprint instead of weight as the attribute, U.S. regulators state that footprint’s primary drawback is that it is less strongly correlated with fuel economy than is weight.<sup>13</sup> Thus, attribute basing, deployed optimally, may offer significant improvements over a flat standard via tagging, but we see little reason to believe this has motivated real policies or that actual policies create significant benefits related to tagging.<sup>14</sup>

We reach a similar conclusion regarding the energy efficiency gap, which is the theory that markets underprovide energy efficiency due to consumer undervaluation or some other market failure (see Jaffe and Stavins (1994)). The energy efficiency gap is frequently cited as a reason for policy intervention, but we are skeptical that it could rationalize attribute basing as it is practiced. If all consumers undervalued energy efficiency by the same amount, then a flat standard with compliance trading can still achieve the first best. If undervaluation was heterogeneous and correlated with  $a$ , then, just as in the second-best targeting case described above, a policy that made  $s$  vary as a function of  $a$  would be very beneficial. But, linear ABR are poorly designed for this purpose. In addition, we are aware of no evidence that this has been considered a rationale for attribute basing in existing policies, and recent evidence has suggested that there is not substantial undervaluation in the automobile market.<sup>15</sup>

### 3 Identifying Attribute Distortions through Bunching Analysis

Our theory indicates that the pivotal determinant of the costs of attribute basing is the elasticity of the secondary attribute with respect to policy incentives. If the attribute does not change at all in response to a policy, then attribute basing will create no efficiency costs. But, a large response of the attribute signifies greater deadweight loss, as described in Proposition 3. The first phase of our empirical analysis is to establish whether attribute basing does indeed distort the choice of the attribute.

We do so by analyzing the distribution of the secondary attribute for the case of Japanese fuel-economy standards. The Japanese regulation has several advantages from the point of view of iden-

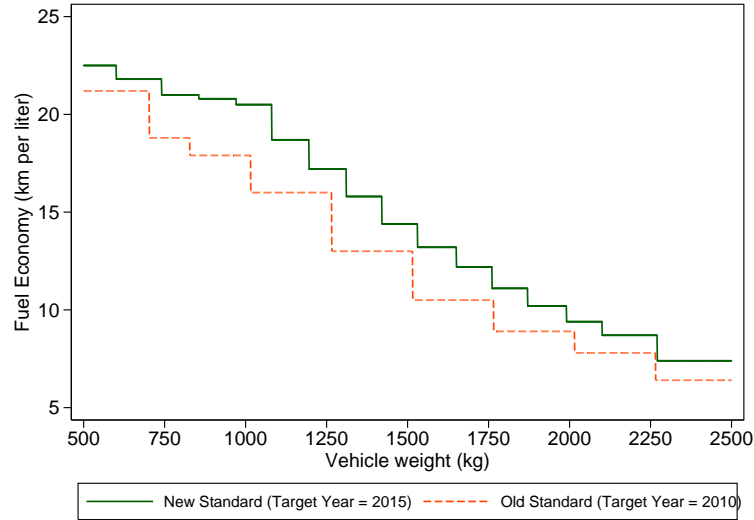
---

<sup>13</sup>See the Federal Register, volume 77, number 199, page 62687. They (sensibly) cite the fact that footprint might be harder to manipulate as its principal benefit.

<sup>14</sup>Jacobsen, Knittel, Sallee, and van Benthem (2015) demonstrate that a fuel economy policy that could regulate based on both fuel economy ( $e$ ) and product durability ( $a$ ) could improve greatly over a policy based only on fuel economy. The structure of such a policy would differ fundamentally from the linear attribute-based policies observed in reality. This is an example of how an optimally designed ABR might have significant welfare implications, but such a policy would not at all resemble the policies we observe.

<sup>15</sup>See Allcott and Greenstone (2012) for a review and Busse, Knittel, and Zettelmeyer (2013); Allcott and Wozny (2014); Sallee, West, and Fan (2016) and Grigolon, Reynaert, and Verboven (2014) for evidence.

**Figure 3:** Fuel-Economy Standards 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.

tification. First, the regulation features “notches”; the fuel-economy target function in Japan is a downward-sloping *step function* in vehicle weight. These notches provide substantial variation in regulatory incentives and allow us to use empirical methods developed for the study of nonlinear taxation (Saez 1999, 2010; Chetty, Friedman, Olsen, and Pistaferri 2011; Kleven and Waseem 2013). Second, the Japanese government has been using attribute-based regulation for decades, and we have more than ten years of data available for analysis.<sup>16</sup> Third, our data span a policy change that enables us to use panel variation, which we pursue in section 4.

### 3.1 Data and Policy Background

Japanese fuel-economy standards, which were first introduced in 1979 and have changed four times since, are weight-based. Our data, which begin in 2001, span the two most recent policy regimes. The target functions for these policies are shown in Figure 3. To be in compliance with the regulation, firms must have a sales-weighted average fuel economy that exceeds the sales-weighted average target of their vehicles, given their weights, and there is no trading of compliance credits across firms although it is one of the key issues in the ongoing policy debate.<sup>17</sup> In Japan, firms are required to meet this

<sup>16</sup>In contrast, the United States just recently instituted attribute-based fuel-economy regulation in 2012, which generates little data for analysis. In addition, its attribute-based target function is smooth, making identification more challenging. Note that the distinction between light-trucks and cars in the original CAFE program is a coarse form of attribute-basing.

<sup>17</sup>Technically, this obligation extends to each weight segment separately. However, firms were allowed to apply excess credits from one weight category to offset a deficit in another. Thus, in the end, the policy is functionally equivalent to

standard only in the “target year” of a policy. This is different from the U.S. CAFE program, which requires compliance annually.<sup>18</sup>

When introducing a new standard, the Japanese government selects a set of weight categories (the widths of the steps in Figure 3).<sup>19</sup> The height of the standard is then determined by what is called the “front-runner” system. For each weight category, the new standard is set as a percentage improvement over the highest fuel-economy vehicle (excluding vehicles with alternative power trains) currently sold in that segment. When the newest standard was introduced in 2009, the government also introduced a separate tax incentive that applies to each specific car model, rather than for a corporate fleet average. We make use of this policy in our panel analysis in section 4. For our bunching analysis, we simply note that both incentives are present in the latter period, and either could be motivating strategic bunching of vehicle weight at the regulatory thresholds (which are common across the two policies).<sup>20</sup>

Our data, which cover all new vehicles sold in Japan from 2001 through 2013, come from the Japanese Ministry of Land, Infrastructure, Transportation, and Tourism (MLIT). The data include 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. Table 1 presents summary statistics. 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. The data are not sales-weighted; we use the vehicle model as our unit of analysis throughout the paper.<sup>21</sup>

---

the U.S. CAFE program, where there is one firm-wide requirement (but no trading across firms).

<sup>18</sup>This does not mean, however, that firms have no incentive to comply before the target year. Consumers see a car’s fuel economy relative to the standard when buying a car, so compliance may affect sales. Compliance may also be a part of long-run interactions between firms and the government. Our data show clearly that firms react to the standards even before the target year. Also, to be precise, under CAFE firms may do some limited banking and borrowing, so they must meet the standard every year, on average.

<sup>19</sup>It is not transparent how these weight categories are chosen, but note that they are almost all of regular width, either 250 kilograms in the old standard or 120 kilograms in the new.

<sup>20</sup>The front-runner system creates strategic incentives for firms to potentially push up the standard further if this might benefit their product line relative to their competitors. We believe that these concerns are unlikely to have an important impact on our analysis, as it affects only a handful of products that are potentially front runners, and in that most of our analysis is focused on weight, not fuel consumption.

<sup>21</sup>Sales-weighting might be a useful extension for some of our results, but Japanese automobile sales data suffer from a problem common to automotive sales data sets in general, which is that sales data are generally recorded at a notably higher unit of analysis and a different calendar. For example, there will be several different versions of the Toyota Camry recorded in our regulatory data, but industry sources typically record sales only for all versions of the Camry together. In addition, the relevant sales are model year totals, not calendar totals, whereas industry data typically cover calendar time and do not distinguish between, for example, a 2013 Camry and a 2014 Camry that are sold in the same month. In contrast, the dataset used in our analysis provides disaggregated data for each vehicle configuration. For example, our dataset provides information about each version of the 2013 Camry as well as each version of the 2014 Camry.

**Table 1:** 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.

### 3.2 Excess Bunching at Weight Notches

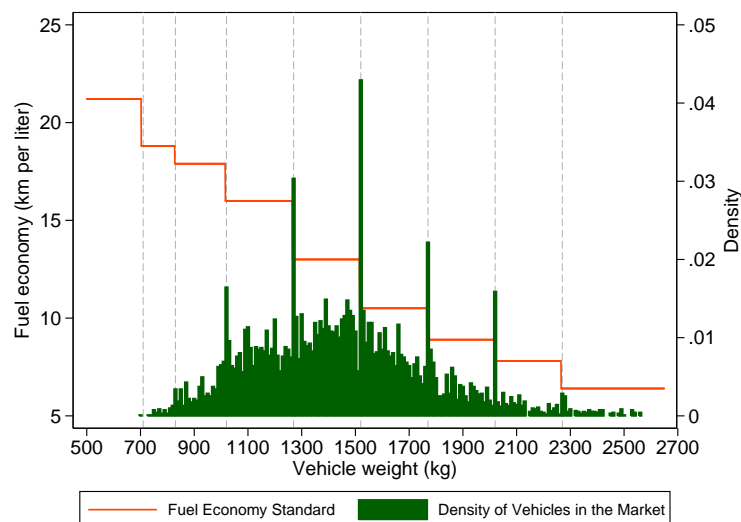
The notched attribute-based standards in Japan create incentives for automakers to increase vehicle weight, but only up to specific values. Increasing weight offers no regulatory benefit, unless the increase passes a vehicle over a threshold. Excess mass (“bunching”) in the weight distribution at exactly (or slightly beyond) these thresholds is thus evidence of weight manipulation. Moreover, if automakers are able to choose vehicle weight with precision, then all manipulated vehicles will have a weight exactly at a threshold. In turn, this implies that all vehicles with weights not at a regulatory threshold have not had their weight manipulated in response to the policy.

We begin by presenting histograms of vehicle weight in Figure 4. Panel A shows the histogram of cars sold between 2001 and 2008. In this period, all vehicles were subject to the old fuel-economy standards, which is overlaid in the same figure. The figure reveals strong evidence of weight manipulation in response to the policy; there is visible excess mass at the notch points. The magnitude of this bunching is substantial. There are more than double the number of cars at each notch point, as compared to the surrounding weight segments. The histogram bins have a width of only 10 kilograms (which is the finest measure available in our data), and yet the bunching appears to be isolated to the weight categories immediately at each threshold. This suggests that automakers can manipulate weight very finely.

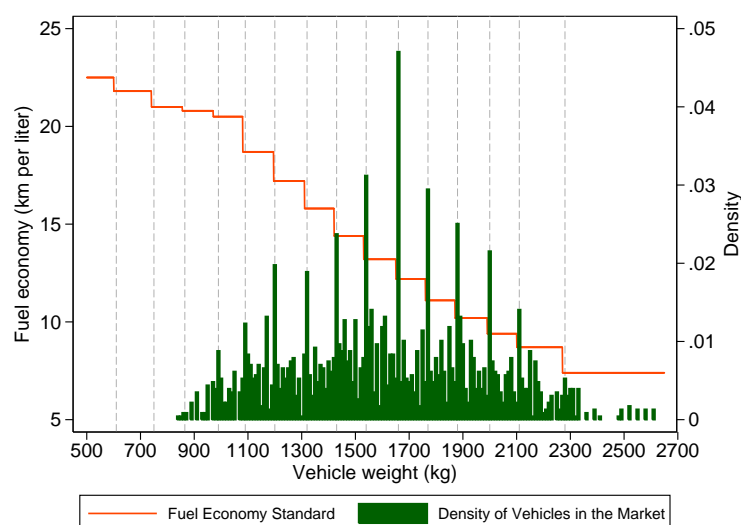
Panel B shows the corresponding figure for data taken from years when the new standard was in effect. Between Panel A and B, the mass points shifted precisely in accordance with the change in

**Figure 4: 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.

the locations of the notch points. This shift in the location of the distributional anomalies provides further compelling evidence that firms respond to the attributed-based regulation.

In the appendix, we present analogous results for “kei-cars”. Kei-cars are very small cars with engine displacements below 0.66 liters.<sup>22</sup> The weight distribution of kei-cars also exhibits bunching at

<sup>22</sup>For comparison, no car sold in the United States in 2010 would qualify as a kei-car. The two-seat Smart Car has the smallest displacement of any car in the United States that year, at 1.0 liters. Although kei-cars must comply with the same fuel-economy regulations as all other cars, we present our results for them separately because they occupy a unique

notch points, and the bunching moves over time in accordance with the policy change.

In sum, the raw data provide strong evidence that market actors responded to attribute-based fuel-economy policies in Japan by manipulating vehicle weight, as predicted by theory. In the next section, we use econometric methods to estimate the magnitude of this excess bunching.

### 3.3 Estimation of Excess Bunching at Notches

Econometric estimation of excess bunching in kinked or notched schedules is relatively new in the economics literature.<sup>23</sup> Saez (1999) and Saez (2010) estimate the income elasticity of taxpayers in the United States 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) use 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 cars into small weight bins (10 kg bins in the application below). For bin  $j$ , we denote the number of cars in that bin by  $c_j$  and the car weight by  $w_j$ . For notches  $k = 1, \dots, K$ , we create dummy variables  $d_k$  that equal one if  $j$  is at notch  $k$ . (Note that there are several bins on each “step” between notches, which we can denote as  $j \in (k - 1, k)$ .) 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:<sup>24</sup>

$$c_j = \sum_{s=0}^S \beta_s^0 \cdot (w_j)^s + \sum_{k=1}^K \gamma_k^0 \cdot d_k + \varepsilon_j, \quad (5)$$

where  $\beta_s^0$  is an initial estimate for the polynomial fit, and  $\gamma_k^0$  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 market segment, have different tax and insurance regulations, and are generally viewed as a distinct product category by Japanese consumers.

<sup>23</sup>See Slemrod (2010) for a review of this literature.

<sup>24</sup>We use  $S = 7$  for our empirical estimation below. Our estimates are not sensitive to the choice of  $S$  for the range in  $S \in [3, 11]$ .

the notches, defined as the 10 kg category starting at the notch. 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 cars 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 cars 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* car weight—that is, excess bunching should come from the “left”. We assume that this is the case.<sup>25</sup> We also make the conservative assumption that the bunching observed at a given notch comes only from the adjacent step in the regulatory schedule, which limits the maximum increase in weight. That is, the bunching at notch  $k$  comes from bins  $j \in (k - 1, k)$ .<sup>26</sup> In practice, automakers may increase the weight of a car 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.

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:

$$c_j + \sum_{k=1}^K \alpha_{kj} \cdot \hat{B}_k = \sum_{s=0}^S \beta_s \cdot (w_j)^s + \sum_{k=1}^K \gamma_k \cdot d_k + \varepsilon_j, \quad (6)$$

where  $\hat{B}_k = c_k - \hat{c}_k = \hat{d}_k$  is the excess number of cars 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

---

<sup>25</sup>The regulation creates an incentive to *increase* car weight in order to bunch at a weight notch because it provides a lower fuel-economy target. The regulation may also create an incentive to decrease weight if, for example, decreasing weight mechanically helps improving fuel economy. However, such an incentive is “smooth” over any weight levels in the sense that vehicles at anywhere in the weight distribution have this incentive, and therefore, the incentive does not create bunching at the notches.

<sup>26</sup>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 biased towards zero.



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.

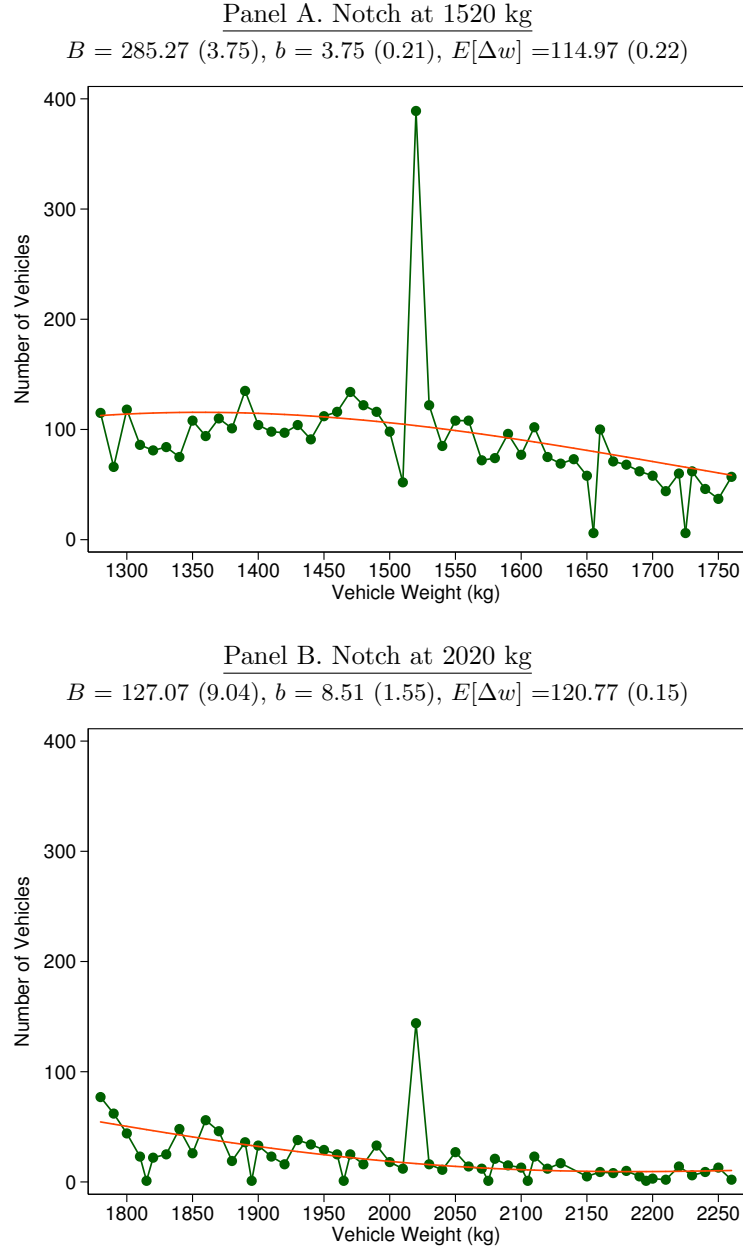
The uniform assumption may underestimate or overestimate  $\Delta w$  if the bunching comes disproportionately from the “left” or the “right” portion of  $j \in (k-1, k)$ . For example, if most of the excess mass comes from the bins near  $k$ , rather than the bins near  $k-1$ , the uniform assumption will overestimate  $\Delta w$ . In practice, this appears to be a minor concern, because the empirical distribution in Figure 4 shows that there are no obvious holes in the distribution, which suggests that the uniform assumption is reasonable. 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. We propose instead an approach that defines  $\alpha_j$  based on the empirical distribution of cars 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  $\alpha_{kj} = \frac{\theta_j}{\sum_{j \in (k-1, k)} \theta_j}$ . In this approach,  $\alpha_{kj}$  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 cars at notch  $k$ ), we provide two more estimates that are relevant to our welfare calculations. The first is the excess bunching as a proportion, which is defined as  $\hat{b} = c_k/\hat{c}_k$ . This is the number of vehicles at a weight notch divided by the counterfactual estimate for that weight. The second estimate is the average changes in weight for cars at notch  $k$ , which is the quantity-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, which follows Chetty, Friedman, Olsen, and Pistaferri (2011) and Kleven and Waseem (2013). We draw from the estimated vector of errors  $\epsilon_j$  in equation (6) with replacement to generate a new set of vehicle counts and follow the steps outlined above to calculate our estimates. We repeat this procedure and define our standard errors as the standard deviation of the distribution of these estimates.

Figure 5 depicts our 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 cars  $B$  is 285.27 (3.75). That is, there are 285 excess cars at this notch compared to the counterfactual distribution.

**Figure 5:** Graphical Illustration of Estimation of Excess Bunching at Each Notch Point



Note: This figure graphically shows the estimation in equation (6). 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.

Bunching as a proportion  $b$  is 3.75 (0.21), which means that the observed distribution has 3.75 times more observations than the counterfactual distribution at this 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 2 presents our estimates for all notches for the data between 2001 and 2008 (the old fuel-

**Table 2:** Excess Bunching and Weight Increases at Each Notch: Old Fuel-Economy Standard

Notch Point	Fuel Economy	Main Estimates			Uniform Assumption		
	Standard below & above the Notch (km/liter)	Excess Bunching	Excess Bunching	$E[\Delta\text{weight}]$	Excess Bunching	Excess Bunching	$E[\Delta\text{weight}]$
		(# of cars)	(ratio)	(kg)	(# of cars)	(ratio)	(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.
<u>Kei-Cars</u>							
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 6. Bootstrapped standard errors are in the parentheses.

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 are few kei-cars in this weight range. Second, we find substantial heterogeneity in the estimates across the notches. The proportional excess bunching  $b$  ranges from 2.1 to 8.5. The estimated weight increases  $E[\Delta w]$  range between 40 kg to 93 kg for

**Table 3:** 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			Uniform Assumption		
		Excess Bunching (# of cars)	Excess Bunching (ratio)	E[ $\Delta$ weight] (kg)	Excess Bunching (# of cars)	Excess Bunching (ratio)	E[ $\Delta$ weight] (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.
<b>Kei-Cars</b>							
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 6. Bootstrapped standard errors are in the parentheses.

kei-cars and 52 kg to 147 kg for other cars. For most cars, this amounts to around a 10% increase in weight, which is substantial. Third, our two different approaches for approximating the counterfactual distribution (uniform or not) produce broadly similar results. Our estimates for  $B$  and  $b$  are not sensitive to the uniform assumption because the excess bunching is very large compared to the counterfactual distribution, so that the way that we reach the integration constraint matters little. We find

slightly larger differences in  $E[\Delta w]$  between our two methods. With the uniform assumption,  $E[\Delta w]$  equals half the width of the regulatory weight step immediately below the notch by assumption, and therefore we have no standard errors for them. Our main estimates do not impose this assumption, but nevertheless yield similar results.

Note that the counterfactual distribution we estimate represents the distribution of vehicle weights that would exist if there was a flat (not attribute-based) fuel-economy standard with the same shadow price.<sup>27</sup> To see this, consider a policy with compliance trading and two weight categories, and thus one notch, at weight  $\tilde{a}$ . The shadow price term in the consumer’s optimization problem is equal to  $\lambda \cdot (e_n - \kappa_1)$  for  $a < \tilde{a}$  and  $\lambda \cdot (e_n - \kappa_2)$  for  $a \geq \tilde{a}$ , where  $\kappa_1 > \kappa_2$ . Thus, for *any*  $a$ , the shadow price term is equal to  $\lambda$  times  $e_n$  minus some constant. The marginal regulatory incentive affecting the choice of  $e$  is thus the shadow price  $\lambda$ , regardless of the weight category. The marginal incentive affecting the choice of  $a$  is zero because a small change in  $a$  does not affect the shadow price term, unless  $a = \tilde{a}$ , in which case there is a discrete jump in the regulatory incentive. Thus, the distortions in  $a$  under the notched policy comes only from the vehicles that bunch at a weight threshold, and the marginal incentives for  $e$  ( $=\lambda$ ) and  $a$  ( $=0$  away from the notches) match those in a flat standard with the same shadow price  $\lambda$ .<sup>28</sup>

Table 3 presents corresponding estimates for all notches for the data between 2009 and 2013 (the new fuel-economy standard period). Note that the new fuel-economy standard has more, and narrower, notches. This mechanically lowers our estimates for  $E[\Delta w]$ , because our conservative approach assumes that automakers do not increase weight to move more than one step.<sup>29</sup> Results from the second policy follow the same pattern. 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 cars 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

<sup>27</sup>This is not the same counterfactual as one with no policy at all. Firms may respond to a flat policy by downsizing vehicles (lowering weight) as part of a strategy to boost fuel economy.

<sup>28</sup>The Japanese policy does not have full compliance trading across firms, so the precise statement is that the counterfactual represents the distribution of weight under a flat subsidy in which each firm faces the same shadow price as in the actual policy. If there were optimization frictions, as in Chetty et al. (2011), then we might expect that some of the vehicles located discretely above the thresholds are also bunching. Our raw data suggest that automakers are able to manipulate weight precisely, because we see excess mass right at each threshold, which suggests that this is not an important issue in our context.

<sup>29</sup>Our panel data do suggest that some weight changes are large enough to cover two steps, but we have no grounds for asserting what fraction of vehicles have been thus altered in our cross-sectional analysis, leading us to prefer providing a reliable lower bound on weight changes.

assumption provides similar estimates to our main estimates.<sup>30</sup>

Overall, the results in this section provide evidence that automakers respond to the attribute-based fuel-economy regulation by changing the weight of vehicles. We find that about 10 percent of vehicles in the Japanese car market manipulated their weight to bunch at regulatory notch points. For these vehicles, the average weight increase induced by the regulation is 110 kg, which is about a 10 percent increase in vehicle weight. This weight increase has welfare implications, as described in our theory above, but it also has implications for safety-related externalities. Heavier vehicles are more dangerous to non-occupants, and when this unpriced, the weight distortions we document here exacerbate safety externalities. We briefly discuss this issue in the next subsection before moving on to our panel analysis.

### 3.4 Safety-related welfare implications of weight manipulation

Our theory emphasized that attribute basing creates a welfare loss because it causes a distortion in the choice of the attribute. The welfare implications of attribute manipulation are altered when the attribute is related to another unpriced externality because the existence of a second externality implies that the privately optimal choice of the attribute is not socially optimal.<sup>31</sup>

In the event of a traffic accident, heavier automobiles are safer for the occupants of the vehicle (this is a private benefit) but more dangerous for pedestrians or the occupants of other vehicles (this is an externality).<sup>32</sup> Thus, the optimal attribute-based policy should *tax* vehicle size rather than subsidize it. The implicit subsidy on vehicle weight in the Japanese fuel-economy standards therefore exacerbate accident-related externalities.

We obtain a back-of-the-envelope estimate of the magnitude of this distortion by multiplying our estimate of the average change in car weight by an estimate of the increased probability that a heavier vehicle causes a fatality during its lifetime, times an estimate of the value of a statistical life. Specifically, the weighted average increase across all cars that bunch at the notches in Table 2 is 109.62 kg. Anderson and Auffhammer (2014) estimate that an increase in vehicle weight of 1000 pounds (454

---

<sup>30</sup>As mentioned above, the new fuel-economy standards were introduced with a separate subsidy incentive that applied to each specific car model. Therefore, the bunching in the new fuel-economy schedule may come from the incentives created by either policy. The bunching in the old fuel-economy schedule comes only from the incentives created by the fuel-economy standards because there was no separate subsidy incentive. We analyze the new policy's subsidy incentive in section 4.

<sup>31</sup>It is straightforward to incorporate a second externality into our framework. In the simplest case when  $a$  causes a separate externality, the optimal attribute-slope will be designed to create a Pigouvian tax on  $a$ . Attribute basing simply provides a second policy instrument, which is necessary for dealing with a second market failure.

<sup>32</sup>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 covers 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.

kg) is associated with a 0.09 percentage point increase in the probability that the vehicle is associated with a fatality, compared to a mean probability of 0.19 percent. For the value of a statistical life, we use \$9.3 million, which comes from a study in Japan (Kniesner and Leeth 1991), and is within the range of standard estimates from the United States.

Multiplying through, we calculate the welfare loss, per car sold, for a 110 kg weight increase as:  $110 * 0.0009 * (2.2/1,000) * \$9.3 \text{ million} = \$2026$  per car that changes weight in response to the policy. Our analysis suggests that the excess bunching accounts for about 10 percent of cars in the market. The Japanese car market sells around 5 million new cars per year, so we estimate our aggregate annual welfare distortion to be 10 percent of 5,000,000 times \$2026, which is \$1.0 billion. For context, automaker revenue in Japan is roughly \$150 billion per year, and Toyota's global revenue in 2013 was \$210 billion. Equivalently, our calculation implies that each new cohort of 5 million cars sold in Japan will be associated with an extra 103 deaths over the lifetime of those cars, which compares to an annual fatality rate of roughly 6,000. Our calculations are meant only as back-of-the-envelope estimates, but they make clear that the welfare distortions induced by the Japanese policy are economically significant.

## 4 Comparing Costs and Benefits of Attribute Basing

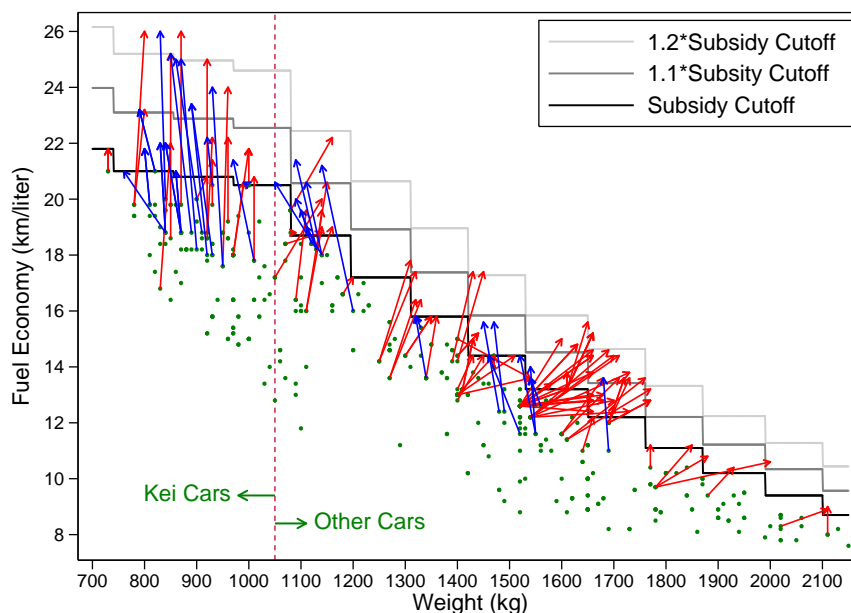
Our theory emphasizes that attribute basing creates welfare costs by distorting the choice of the secondary attribute. The bunching analysis in section 3 provides empirical evidence for this prediction. But, our theory also shows that, when each product must individually comply with a standard (i.e., when there is no compliance trading), an ABR may have efficiency benefits stemming from marginal cost equalization. However, even when an ABR is superior to a flat standard, it is likely to be inferior to a flat standard with trading.

The relative efficiency of the three policies (an ABR, a flat standard, and a flat standard with trading) is an empirical question because it depends on the elasticity of the attribute as well as the degree of marginal cost equalization achieved by the ABR. Moreover, our theory suggests that ABR might provide distributional benefits even if it is inefficient. To provide empirical evidence for these predictions from our theory, we develop an empirical strategy that exploits quasi-experimental variation created by a subsidy policy in Japan.

### 4.1 A Subsidy Policy and Descriptive Evidence from Panel Data

In 2009, the Japanese government introduced a new subsidy that applied to each car model rather than the corporate average. In Figure 6, we visualize this subsidy policy. If a car had fuel economy

**Figure 6:** The Subsidy’s Eligibility Cutoffs and the Revealed Responses to the Incentive



Note: The solid lines are three step functions that correspond to the three tiers of the new subsidy’s eligibility cutoffs in 2012. A car had to be above the eligibility cutoff line to obtain the subsidy. The scatterplot shows each car’s fuel economy and weight in 2008—the year before the introduction of the subsidy. For the cars that qualified for the subsidy in 2012, we also show “arrows” connecting each car’s starting values in 2008 with its values in 2012.

higher than the subsidy cutoff—the lowest step-function in the figure, which is equivalent to the new fuel-economy standard presented in Figure 3—consumers purchasing that car received a direct subsidy of approximately \$700 for kei-cars and \$1,000 for other cars.<sup>33</sup> In addition, cars with fuel economy 10% and 20% higher than the subsidy cutoff received more generous subsidies in the form of tax exemptions. This creates what we call a “double notched” policy—a car had to be above a step-function in the two-dimensional space of fuel economy and weight.

This policy provides two advantages in studying the costs and benefits of attribute basing. First, it created quasi-experimental variation in incentives to change weight ( $a$ ) and fuel economy ( $e$ ). Even though all products faced the same subsidy cutoff, each had a different set of *changes* in  $a$  and  $e$  that would be required to get the subsidy. This variation comes from differences in starting points, from the introduction of new weight notches, and from the fact that the changes in the standards are different across the weight categories. As a result, some vehicles are able to make modest improvements in fuel economy to gain the subsidy, whereas others require large changes. And, some vehicles can take advantage of a weight notch with small increases in weight, but others require a large increase. These

<sup>33</sup>This policy was called the “eco-car subsidy.” The government introduced the policy in April, 2009. The policy was effective in 2009, parts of 2010 and 2011, and 2012. In 2012, the subsidy was 100,000 JPY (approximately 1,000 USD using the exchange rate in 2012) for all passenger cars except for kei-cars, which received 70,000 JPY (approximately 700 USD).



differences create a rich source of identification for our estimation.

Second, our theory suggests that when compliance trading is not available, attribute basing creates a trade-off between distortionary costs and efficiency benefits. The model-specific subsidy indeed creates such an environment because each individual car must comply to gain the subsidy. That is, attribute basing may play a role in improving efficiency in this subsidy policy, while such benefits are likely be minimal for the fleet-average regulation.

We begin by describing raw data that reveal how products changed in response to the subsidy. We construct panel data of 675 cars by linking cars sold in 2008 (before the policy change) and 2012 (the last year of our data) based on a unique product identifier (ID) in the regulatory data.<sup>34</sup> Each dot in Figure 6 shows a car’s starting values of fuel economy and weight in 2008. For the cars that qualified for the new subsidy in 2012, we also show vectors connecting each car’s starting position in 2008 to its final position in 2012. This figure is an empirical analog to Figure 2 in our theory section.

The raw data reveal several useful results. First, many of the cars that gained the subsidy were redesigned in a way that they were just above the subsidy cutoff. Second, we see that some cars had large weight increases, which is consistent with our findings in the bunching analysis. Third, most cars except for kei-cars increased both weight and fuel economy in order to gain the subsidy—that is, they moved “northeast” in the figure. Fourth, the attribute schedule is nearly flat in the weight distribution inhabited by kei-cars, so they are subject to a nearly-flat (not attribute-based) policy.<sup>35</sup> Most kei-cars that obtained the subsidy increased fuel economy, while slightly reducing weight. This notable difference between kei-cars and other cars accords with our theory: the steeper sloped portion of the standard induced substantial weight manipulation, while the flatter sloped part of the standard did not. Finally, cars that started off closer to the new standard were more likely to get the subsidy; that is, the “distance” to the subsidy cutoff explains most of the variation in which cars obtained the subsidy.<sup>36</sup>

## 4.2 A Discrete Choice Model of Vehicle Redesign

To translate the revealed preference data into useful information, we develop a discrete choice model based on the theoretical framework in section 2. Suppose that we observe data from a competitive market in two time periods, the first of which has no policy, and the second of which features the model-

---

<sup>34</sup>We describe how we construct the panel dataset in the Appendix D.

<sup>35</sup>The figure shows an approximate dividing line between kei-cars and other cars. Kei-cars are not regulated by weight, but rather by engine displacement, so this division is not strict.

<sup>36</sup>In the appendix, we also show the corresponding arrows for vehicles that did not receive a subsidy. Most of these vehicles had much smaller changes in fuel economy and weight, which suggests that secular trends were modest during this period. We also find that some of these cars increased weight to be at the weight notch underneath the subsidy cutoff lines. This is because, in addition to the model-specific subsidy incentive, vehicles were still influenced by the corporate average fuel-economy standards, which gave them an incentive to increase weight. Our estimation method below controls for the effects of the corporate average fuel-economy standards.

specific subsidy. We follow our theory in section 2 to consider a consumer type  $n$  who purchases a car with weight ( $a_n$ ) and fuel economy ( $e_n$ ). Suppose that  $(a_n^0, e_n^0)$  is the pair of fuel economy and weight in the first period. This pair should be welfare maximizing given the firm's production function and the tastes of its consumers. Any deviation from this bundle would generate private welfare loss  $L(\Delta a_n, \Delta e_n)$  as presented in Figure 2 in the theory section.

When the subsidy policy is introduced, this product either stays at the initial optimal pair  $(a_n^0, e_n^0)$  or moves to a new pair of  $(a_n, e_n)$  that makes the product eligible for the subsidy at the lowest possible loss. Figure 6 provides two pieces of information relevant to this optimization problem. First, the loss  $L(\Delta a_n, \Delta e_n)$  was likely to be larger than the subsidy's value for the cars that did not move to the subsidy cutoff. Second, for the cars that moved above the subsidy cutoff, the arrows reveal their optimal paths to comply with the requirement, which provides information about the loss function.

Based on this logic, we develop a discrete choice problem of vehicle redesign. With the introduction of the subsidy, each product  $n$  chooses the new optimum pair of  $(a_n, e_n)$ . Following our notations in the theory section, we use  $\Delta a_n = a_n - a_n^o$  and  $\Delta e_n = e_n - e_n^o$  to denote product  $n$ 's deviation from its initial optimum, which generates the private welfare loss  $L(\Delta a_n, \Delta e_n)$ .

We make two important notes on the interpretation of  $L(\Delta a_n, \Delta e_n)$ . First, this private welfare loss is generated by a regulation. A car's pair of  $(a_n, e_n)$  is changed to comply with a regulatory incentive, which is the subsidy in this case. For this reason, we can interpret  $L(\Delta a_n, \Delta e_n)$  as a regulatory compliance cost. Second, we emphasize that  $L(\Delta a_n, \Delta e_n)$  does not identify primitives of either the utility function or the cost function, but instead provides a sufficient statistic that reveals the change in welfare due to the introduction of the subsidy policy. Note that if consumer types match perfectly to a unique vehicle and the market is perfectly competitive, then  $L(\Delta a_n, \Delta e_n)$  can be interpreted directly as a change in consumer surplus. Allowing for firms to price above marginal cost does not necessarily alter this interpretation.<sup>37</sup> If a markup for  $n$  does not change before and after the introduction of the subsidy, the loss function still represents changes in consumer surplus.<sup>38</sup> When markups change, then  $L(\Delta a_n, \Delta e_n)$  may over or understate the impact on consumers depending on how markups change.

We describe the choice of the new optimum  $(a_n, e_n)$  for product  $n$  as the outcome of a discrete

---

<sup>37</sup>If we maintain the assumption that consumer types are matched to a vehicle, then firms will invest in  $a$  and  $e$  up until the marginal cost of doing so meets the marginal benefit of the consumer, as this will allow them to raise prices accordingly (see Reynaert and Sallee (2016) for a closely related model).

<sup>38</sup>It is plausible that markups did not change much between the two years in our setting. The logic of differentiated product market equilibrium under Bertrand competition, as exemplified in Berry, Levinsohn, and Pakes (1995), says that the markup for a vehicle, in equilibrium, is determined by its market share, the price elasticity of demand, and the first order condition from the Bertrand competition. We calculate each vehicle's market share in 2008 and 2012 using sales data. We find that changes in market share are very small. The 25th, 50th, and 75th percentiles of the change in market share are  $-0.0004$ ,  $-0.0001$ , and  $0.0002$ . For the range of reasonable price elasticities, including estimates from the literature on the Japanese and U.S. automobile markets, these small changes in market shares would imply economically insignificant changes in optimal markups.

choice over all of the possible (discretized) grid points in  $a$  by  $e$  space.<sup>39</sup> We denote each grid point as a unique value of  $z$ . The second-period optimization problem for product  $n$  is to choose  $(a_n, e_n)$  that maximizes  $W_{nz}$ , which is the compliance cost plus the subsidy:

$$W_{nz} = L(\Delta a_n, \Delta e_n) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz}. \quad (7)$$

Consistent with our theoretical model, we assume that the loss function is quadratic. However, we allow the parameters to flexibly vary by  $n$  through random-coefficients so that  $L(\Delta a_n, \Delta e_n) = \alpha_n(\Delta a_n)^2 + \beta_n(\Delta e_n)^2 + \gamma_n \Delta a_n \Delta e_n$ .  $\tau$  is also a parameter to be estimated, which represents the value of receiving the subsidy, and  $\varepsilon_{nz}$  is a Type-I extreme value error term specific to each vehicle  $n$  and each grid point  $z$ . First, we estimate equation 7 via a logit. Second, we relax the logit assumptions by allowing parameters to vary by product  $n$  via a random-coefficients logit. When we interpret the parameters in terms of dollars, we rescale them by the dollar value of the subsidy by dividing by  $\hat{\tau}$  (Train 2009).

Table 4 presents estimates from the logit model.  $\Delta$ Weight is a change in weight in 100 kilogram units and  $\Delta$ (Fuel consumption) is a change in liters per 100km (l/100km).<sup>40</sup> The coefficient on the interaction term is positive, which implies that the optimal path of reducing fuel consumption (improving fuel economy) should involve a *reduction* in weight. The first two coefficients ( $-0.94$  and  $-1.12$ ) are roughly the same, indicating that a change in weight by 100 kg and a change in fuel consumption by one l/100km result in approximately the same loss.

*Accounting for fleet-average regulation*—The simple specification in equation (7) provides transparent interpretations on the estimates of the loss function. However, it considers only the product-specific subsidy and ignores the fleet-average regulation. In the appendix, we show that the following estimating equation provides estimates of the parameters necessary to identify the loss function, taking into account the effects of the regulations:

$$V_{nz} = L(\Delta a_n, \Delta e_n) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \lambda[e_n - \sigma(a_n)] - \lambda^o[e_n - \sigma^o(a_n)] + \varepsilon_{nz}, \quad (8)$$

where  $L(\Delta a_n, \Delta e_n) = \alpha_n(\Delta a_n)^2 + \beta_n(\Delta e_n)^2 + \gamma_n \Delta a_n \Delta e_n$ . Compared to equation (7), this equation includes two control variables:  $e_n - \sigma(a_n)$  is the position of the vehicle vis-à-vis the new policy and  $e_n - \sigma^o(a_n)$  is the position of the vehicle vis-à-vis the old policy evaluated at the *new* choice of  $a$  and  $e$ . In the appendix, we show that i)  $\lambda$  and  $\lambda_o$  represent the shadow prices of the original and new

<sup>39</sup>While weight and fuel economy are in principle continuous measures, the regulatory data are measured in discrete units (10 kilograms for weight and tenths of a kilometer-per-liter for fuel economy), and all regulations are based on these discrete units.

<sup>40</sup>We use fuel consumption (l/100km) rather than fuel economy (km/liter) here because when regulators calculate the corporate average fuel economy, they average each model's fuel consumption rather than fuel economy (equivalently, they harmonically average fuel economy).

**Table 4:** Estimates of the Loss Function

	Logit		Random-coefficient logit		
	(1)	(2)	(3)	(4)	(5)
$\alpha : (\Delta\text{Weight})^2$	-0.94 (0.08)	-0.97 (0.08)	-1.58 (0.17)	-1.61 (0.17)	-1.61 (0.18)
$\beta : (\Delta\text{Fuel consumption})^2$	-1.12 (0.07)	-1.16 (0.08)	-1.51 (0.12)	-1.55 (0.13)	-1.66 (0.18)
$\gamma : \Delta\text{Weight} \times \Delta\text{Fuel consumption}$	0.26 (0.11)	0.31 (0.11)	0.24 (0.12)	0.30 (0.12)	0.60 (0.17)
$\tau : 1\{\text{Subsidy}\}$	0.66 (0.15)	0.38 (0.18)	0.79 (0.16)	0.50 (0.19)	0.40 (0.20)
$\lambda$ : Shadow price of new fuel-econ. standard		0.42 (0.14)		0.45 (0.15)	0.66 (0.18)
$\dot{\lambda}$ : Shadow price of old fuel-econ. standard		0.18 (0.12)		0.19 (0.12)	0.19 (0.12)
Stadard deviation of random-coefficient					
$\alpha : (\Delta\text{Weight})^2$			0.63 (0.09)	0.63 (0.09)	0.69 (0.10)
$\beta : (\Delta\text{Fuel consumption})^2$			0.48 (0.05)	0.47 (0.05)	0.34 (0.05)
$\gamma : \Delta\text{Weight} \times \Delta\text{Fuel consumption}$			0.91 (0.12)	0.93 (0.12)	0.50 (0.11)
$\lambda$ : Shadow price of new fuel-econ. standard					1.06 (0.28)
$\dot{\lambda}$ : Shadow price of old fuel-econ. standard					0.01 (0.00)

Note: This table shows the estimation results of the discrete choice models in equation (7). We define  $\Delta\text{Weight}$  as each vehicle's change in weight in 100 kilogram units and  $\Delta(\text{Fuel consumption})$  as its change in liters per 100km (l/100km). The estimation is based on the panel data of 675 matched vehicle models.

flee-average regulations, and ii) the loss function excluding the effect of the fleet-average regulation effects is  $\hat{L}(\Delta a_n, \Delta e_n) = L(\Delta a_n, \Delta e_n) + \lambda^o[e_n^o - \sigma^o(a_n^o) - (e_n - \sigma^o(a_n))]$ .

Columns 1 and 2 of Table 4 show that the addition of the regulatory control variables have little impact on the coefficients on the loss function. The subsidy coefficient, however, becomes smaller in this specification. This is because we are likely to overestimate the subsidy effect in columns 1 and 2 since we do not control for the fleet-average regulation. Importantly, while the coefficient on the subsidy affects the dollar values of the loss, it does not have major impacts on the *relative* costs of changing  $e$  versus  $a$  because the estimates on  $(\Delta a)^2$  and  $(\Delta e)^2$  are robust between the two approaches.<sup>41</sup>

*Random-coefficients models*—In columns 3 to 5 of Table 4, we relax the logit assumptions by allowing random coefficients. In the bottom of the table, we report the standard deviations of normally distributed random coefficients. The mean coefficients are similar to those in the logit estimation except that the coefficients for  $(\Delta a)^2$  and  $(\Delta e)^2$  are slightly larger in absolute value in the random-coefficients model. The standard deviations for the random-coefficients are statistically significant and economically meaningful, suggesting that there is substantial heterogeneity in the parameters for the loss function. In our counterfactual policy analysis, we report results based on both the logit and the random-coefficients logit to examine the policy implication of this heterogeneity.

*Predicting bunching from the model*—Before we proceed to counterfactual policy simulation, we ask whether our discrete choice model is able to predict the excess mass of vehicles located around weight thresholds observed in the data we estimated in section 3. This serves as a check on the loss function estimates; it asks whether our panel estimates can predict the excess mass observed in our cross-sectional analysis.

We use the estimated loss function in Table 4 and data from 2008 to predict the amount of bunching in weight in 2012. Using the functional form assumption for  $L(\Delta a_n, \Delta e_n)$ , the estimated coefficients, and data from 2008  $(a_n^o, e_n^o)$ , we calculate the predicted value for the value function  $V_{nz}$  and let each  $n$  choose the optimal pair of  $(a_n, e_n)$  in 2012. We then compare the predicted bunching from this procedure to the actual bunching in the 2012 data.

Results are in Table 5 for the random-coefficients model and Appendix Table A-4 for the logit model. In the data, we observe that 27.6% of cars in 2012 are located at one of the weight notches

---

<sup>41</sup>Though it is not our focus, we do note that this procedure also delivers an estimate of the shadow price of fuel-economy regulations, which are of much interest to the literature. Our procedure differs from the existing literature in leveraging panel data around a policy *change* to identify the shadow price, whereas existing work either examines a specific policy loophole (Anderson and Sallee 2011) or uses static structural models (Goldberg 1998; Gramlich 2009; Jacobsen 2013a). The estimate in column 3 implies a shadow price of \$1,162 (= .45/.37 \* \$1000) per unit of 1/100km car per. This translates into \$258 per mpg per car at the average fuel economy in our sample. Our simple approach does not account for imperfect competition, so we do not stress these results, but simply note that they are the same order of magnitude as results found in Gramlich (2009) and Jacobsen (2013a) for the United States.

**Table 5:** Actual Bunching vs. Predicted Bunching Based on the Estimated Loss Function

	All	860 kg	980 kg	1090 kg	1200 kg	1320 kg	1430 kg
Actual (%)	27.60	1.58	3.17	1.81	0.90	1.58	1.81
Predicted (%)	26.14	1.14	2.27	3.86	2.27	1.36	4.32
	1540 kg	1660 kg	1770 kg	1880 kg	2000 kg	2110 kg	2280 kg
Actual (%)	4.52	5.66	2.71	1.36	1.36	1.13	0.00
Predicted (%)	5.23	1.82	0.91	1.36	0.68	0.91	0.00

Note: We use the estimated loss function obtained by the random-coefficients model and data in 2008 to predict bunching in 2012. The table shows actual and predicted bunching (% of the data at each notch).

in the fuel economy standard. With the logit specification, our model predict that 32.73% of cars bunch at the notches, which suggests that the logit specification does a reasonable job at predicting the bunching. The random-coefficients model predicts that 26.14% of cars bunch at the notches, indicating that allowing heterogeneity helps improve the degree to which the discrete choice model can predict the actual movement of the data.

*Summary*—Our method provides a simple and transparent framework to translate the revealed preference information from the raw data to key empirical parameters for our policy simulation. It establishes an approach for analyzing double notches that may prove useful in other contexts. The price of this simplicity and transparency is that we must make several substantive assumptions. It assumes that the functions determining taste and price are unchanged between the two periods. It is based on only our matched observations, and does not model entry and exit. It interprets the loss function as social cost. This is correct if, as in our exposition, the market is perfectly competitive. Our results would still represent net social cost if the market were imperfectly competitive but markups did not change between the two periods. If markups do change significantly, then our loss function, which should then be interpreted as lost profit, would include a mixture of net social costs and transfers between producers and consumers.

### 4.3 Counterfactual Policy Simulations

Using the estimated loss function, we simulate three counterfactual policies to examine the costs and benefits of attribute basing:

- **ABR:** This is the actual new attribute-based fuel-economy standards in Japan (Figure 3). We use  $\bar{e}$  to denote the improvement in the average fuel economy induced by this policy.
- **Flat:** We construct a flat fuel-economy standard that improves the average fuel economy by  $\bar{e}$ .
- **Efficient:** We construct a flat fuel-economy standard with compliance trading that improves the

average fuel economy by  $\bar{e}$ . Compliance trading equalizes the marginal compliance cost across cars, which makes this policy the most efficient to achieve  $\bar{e}$ .

Letting each policy achieve the same level of an improvement in the average fuel economy ( $\bar{e}$ ) makes the three policies comparable. For these counterfactual policies, we consider an energy-efficiency mandate—each product is required to comply with the standard. We focus on a mandate because our theory suggests that when a policy is implemented as a mandate, attribute basing can be justified by efficiency considerations. However, our theory also suggests that attribute basing is unlikely to be equivalent to the efficient policy unless it can fully equalize the marginal compliance cost. To empirically investigate these questions, we simulate how automakers respond to these policies based on the estimated loss function in Table 4.

*Attribute-based policy*—This policy requires all vehicles in Figure 6 to move above the lowest line. Given each vehicle’s initial point and our estimates of the loss function  $L$ , we find the  $a$  and  $e$  that achieves compliance at the lowest possible cost. We then calculate the resulting  $\Delta a$ ,  $\Delta e$ , and  $\Delta L$  for each car’s new optimal point.<sup>42</sup>

Panel A of Table 6 reports results based on the loss function with logit specification (column 2 of Table 4). The counterfactual ABR lowers fuel consumption  $e$  by 0.76 l/100km on average, which is approximately a 10 percent improvement in fuel economy. Consistent with our theoretical prediction, the ABR induces a 33 kilogram (73.37 lb) average weight increase, which is about a 3 percent increase in weight. This policy’s private compliance cost  $L(\Delta a, \Delta e)$  is \$2272 per unit sold, averaged across all model types. An important statistic for measuring the efficiency of regulation is the standard deviation in the marginal cost of increasing  $e$  at the optimal choice across products. In the most efficient policy, this statistic is zero because the marginal compliance cost is equalized across products. Therefore, the dispersion of the marginal compliance cost provides a measurement of inefficiency. For the ABR, this standard deviation is \$2002. Finally, we calculate the external cost from safety externality in the last column. In addition to the private compliance cost, the ABR creates a cost from the safety externality of \$606 because the policy induces an increase in weight.

*Flat policy*—The flat standard has the benefit of not distorting weight. The estimated  $\gamma$  in Table 4 implies that when firms need to reduce fuel consumption  $e$ , it is helpful to *reduce* weight  $a$ . Consistent with this result, we observe that the flat policy reduces the average weight by 25 kg in Table 6, which is an efficient change in weight rather than a distortion. The flat standard, however, creates many infra-marginal vehicles; that is, vehicles that are in compliance with the standard without any change in fuel economy. These infra-marginal vehicles have a zero marginal cost of increasing  $e$ , but do not

---

<sup>42</sup>In our simulation, we assume that all vehicles stay in the market after the introduction of this simulated policy, although in reality there can be entry and exit.

**Table 6:** Counterfactual Policy Simulations

	$\Delta e$ : Fuel consumption (liter/100km)	$\Delta a$ : Weight (kg)	$L(\Delta a, \Delta e)$ : Compliance cost (\$/car)	Cost relative to efficient	S.D. of marginal compliance cost (\$/car)	External cost from safety externality (\$/car)
Panel A: Based on logit estimates						
Efficient	-0.86	-12.07	1046	1.00	0	-222
Flat	-0.86	-25.31	4038	3.86	4134	-466
ABR	-0.86	32.91	2272	2.17	2002	606
Panel B: Based on random-coefficient logit estimates						
Efficient	-0.86	-5.20	716	1.00	0	-96
Flat	-0.86	-16.38	4637	6.47	4610	-302
ABR	-0.86	36.05	2504	3.50	2158	664

Note: This table shows the results of our three policy simulations: 1) attribute-based fuel-economy standards (ABR), a flat fuel-economy standard with no compliance trading (Flat), and a fully efficient policy that is equivalent to a flat standard with compliance trading (Efficient).

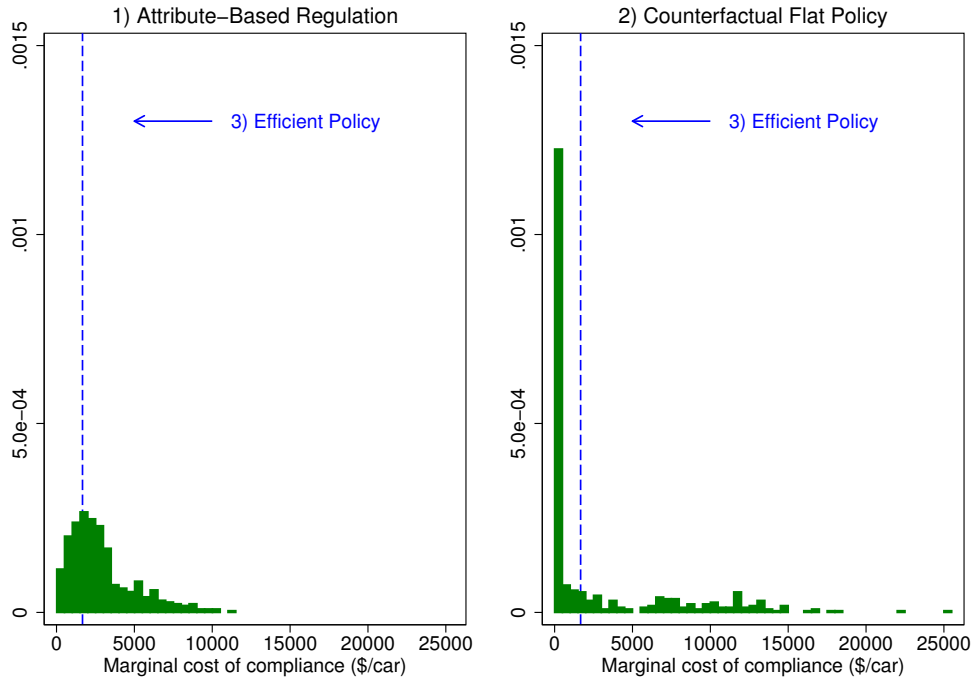
change  $e$  at all, because there is no regulatory incentive. Other vehicles have very large marginal costs of increasing  $e$  because they have to improve fuel economy by a large amount in order to comply with the policy. This dispersion in marginal costs is inefficient, and it results in compliance costs that are, on average, 1.95 times larger than the compliance costs of the ABR. To visualize this inefficiency, Figure 7 plots the distribution of marginal compliance costs across products under the ABR and the flat policy. Under the flat policy, there are many more infra-marginal (zero marginal cost) observations, and the remaining distribution is also more diffuse. The benefit of attribute basing is the (partial) harmonization of these marginal costs.

*Efficient policy*—The ABR’s harmonization in marginal compliance cost is incomplete and that makes attribute basing an inefficient substitute for a compliance trading system. In the efficient policy, firms face a flat standard but are allowed to trade compliance. As shown in our theory, this policy is equivalent to providing the efficient Pigouvian marginal subsidy for an improvement in  $e$ . This policy completely harmonizes marginal costs, no products are infra-marginal, and therefore the total compliance cost is minimized. Because the standard deviation of the marginal compliance cost is zero, its distribution collapses to a constant, which we label in Figure 7. Table 6 confirms that this policy achieves the policy goal by the lowest compliance cost.

*Heterogeneity in compliance cost*—Panel B of Table 6 shows the same statistics using the loss function that allows random-coefficients. To make the results comparable to Panel A, we let each policy produce the same level of  $\bar{e}$  as in Panel A. The results in Panel B are qualitatively the same as



**Figure 7:** Efficiency of Regulation: Dispersions in Marginal Compliance Cost

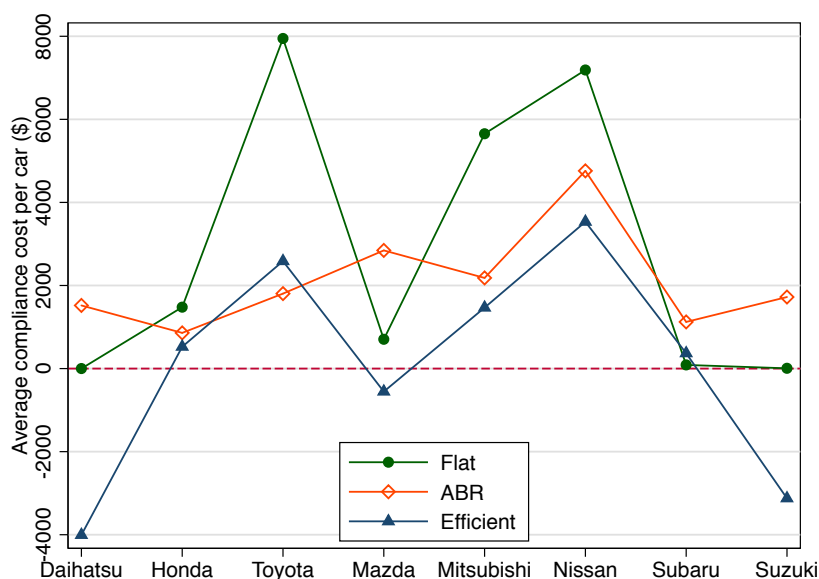


Note: This figure shows the distributions of the marginal cost of compliance for the three simulated policies: 1) the attribute-based regulation (left), 2) the counterfactual flat policy (right), and 3) the counterfactual efficient policy (dashed line). The distribution of the efficient policy is the vertical line because the standard deviation is zero.

in Panel A. The random-coefficients logit differs by introducing heterogeneity in the compliance cost function, which creates an additional disadvantage for ABR relative to the efficient benchmark. This is because the ABR helps harmonize marginal compliance costs only through the correlation between weight and compliance cost, whereas compliance trading harmonizes marginal costs that arise from heterogeneity in the cost function as well. Panel B shows that the ABR performs worse when we use the random-coefficients logit for the simulation.

*Distributional considerations*—Proposition 4 in our theory suggests that attribute basing may exist in real-world policies because of its distributional implications. Although attribute basing is inefficient, it alters the incidence of compliance costs across different firms and consumers, which may be desirable to policy makers. We use our simulation results to provide empirical evidence of the variation in incidence by firm across policy alternatives. In Figure 4.3, we show the average compliance cost per car by firm. A flat standard induces uneven compliance costs across firms partly because some firms produce more heavy cars than others. Attribute basing mitigates this dispersion by allowing heavier cars to improve fuel economy less. As a result, the average compliance cost is more evenly distributed in the ABR. Finally, in the efficient policy, some firms have negative compliance costs because compliance trading allows them to sell excess compliance to others. This figure illustrates the

**Figure 8:** Distributional Consideration: Compliance Cost by Firm



Note: This figure shows each firm’s average compliance cost (\$ per car) for the three counterfactual policies. The cost can be negative in the efficient policy (a flat standard with compliance trading) because firms can sell credits to other firms.

potential efficiency-equity trade-off between the efficient policy and the ABR. Although the efficient policy produces the lowest compliance cost for the society by equalizing marginal compliance cost (Table 6), the ABR might be preferred by policymakers if they wish to prioritize benefits for certain firms and their consumers.

*Summary*—In sum, our counterfactual policy analysis highlights several key policy implications consistent with our theoretical predictions. First, ABR creates substantial distortions in the attribute. An efficient policy leads to a reduction in weight, whereas the ABR causes a noticeable increase in weight. Second, a benefit of ABR is the *partial* harmonization of the marginal costs of compliance. Relative to a counterfactual non-ABR flat policy, ABR produces smaller dispersions of the marginal compliance costs between products. This smaller dispersion produces a smaller total regulatory cost compared to a flat policy. Third, ABR is nevertheless an inefficient substitute for a fully efficient policy such as a non-ABR policy with compliance trading because the efficient policy does not create attribute distortions and fully equalizes the marginal costs of compliance. Fourth, ABR changes the distributional impacts of policy, so that ABR could be used as a way to favor one set of products over others.

## 5 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 improve efficiency by equalizing marginal costs of regulatory compliance, but we emphasize that, even in those cases, the optimal attribute-based policy deviates fundamentally from those observed in real world policies because it must balance efficiency gains from cost equalization with the distortion it induces in the attribute.

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 that take advantage of a “double notched” policy, to demonstrate that the Japanese car market has experienced a notable increase in weight in response to attribute-based regulation. This weight increase exacerbates safety-related externalities to an economically significant degree.

Our analysis of the double notch enables us to make internally consistent comparisons of the benefits from marginal cost equalization and the distortions from increases in weight. Our estimates suggest that, for the Japanese tax policy we consider, attribute basing is significantly more efficient than a flat standard without trading. But, it achieves increases in fuel economy at roughly twice the cost of a flat standard that equalizes compliance costs through a trading system. This suggests that attribute basing can substantially improve welfare when regulations are unable to use market mechanisms to harmonize marginal compliance costs, but attribute basing is nevertheless a second-best solution that deviates significantly from the first-best Pigouvian benchmark.

## References

- Akerlof, George A. 1978. “The Economics of “Tagging” as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning.” *American Economic Review* 68 (1):8–19.
- Allcott, Hunt and Michael Greenstone. 2012. “Is There an Energy Efficiency Gap?” *Journal of Economic Perspectives* 26 (1):3–28.
- Allcott, Hunt and Nathan Wozny. 2014. “Gasoline Prices, Fuel Economy, and the Energy Paradox.” *Review of Economics and Statistics* 96 (5):779–795.
- Anderson, Michael and Maximilian Auffhammer. 2014. “Pounds that Kill: The External Costs of Vehicle Weight.” *Review of Economic Studies* 82 (2):535–571.
- Anderson, Soren T. and James M. Sallee. 2011. “Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards.” *American Economic Review* 101 (4):1375–1409.

- Becker, Randy and Vernon Henderson. 2000. "Effects of Air Quality Regulations on Polluting Industries." *Journal of Political Economy* 108 (2):379–421.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 64 (4):841–890.
- Blinder, Alan S. and Harvey S. Rosen. 1985. "Notches." *American Economic Review* 75 (4):736–747.
- Borenstein, Severin. Forthcoming. "A Microeconomic Framework for Evaluating Energy Efficiency Rebound and Some Implications." *Energy Journal* .
- Brock, William A. and David S. Evans. 1985. "The Economics of Regulatory Tiering." *The Rand Journal of Economics* 16 (3):398–409.
- Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer. 2013. "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review* 103 (1):220–256.
- 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.
- Decarolis, Francesco. 2015. "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?" *American Economic Review* 105 (4):1547–80.
- Dharmapala, Dhammika, Joel Slemrod, and John Douglas Wilson. 2011. "Tax Policy and the Missing Middle: Optimal Tax Remittance with Firm-Level Administrative Costs." *Journal of Public Economics* 95 (9-10):1036–1047.
- Fack, Gabrielle and Camille Landais. 2015. "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France." Manuscript: London School of Economics.
- Gao, Feng, Joanna Shuang Wu, and Jerold Zimmerman. 2009. "Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes-Oxley Act." *Journal of Accounting Research* 47 (2):459–506.
- Gelber, Alexander, Damon Jones, and Daniel W. Sacks. 2015. "Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test." Manuscript: University of California.
- Gillingham, Kenneth. 2013. "The Economics of Fuel Economy Standards versus Feebates." Manuscript: Yale University.
- Gillingham, Kenneth, David Rapson, and Gernot Wagner. Forthcoming. "The Rebound Effect and Energy Efficiency Policy." *Review of Environmental Economics and Policy* .
- Goldberg, Pinelopi Koujianou. 1998. "The Effects of the Corporate Average Fuel Efficiency Standards in the U.S." *The Journal of Industrial Economics* 46 (1):1–33.
- Gramlich, Jacob. 2009. "Gas Prices, Fuel Efficiency, and Endogenous Product Choice in the U.S. Automobile Industry." [Http://faculty.msb.edu/jpg72/Autos\\_GramlichJacob.pdf](http://faculty.msb.edu/jpg72/Autos_GramlichJacob.pdf).
- Grigolon, Laura, Mathias Reynaert, and Frank Verboven. 2014. "Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market." Manuscript: University of Leuven.
- Gruenspecht, Howard K. 1982. "Differentiated Regulation: The Case of Auto Emissions Standards." *American Economic Review* 72 (2):328–331.

- 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.
- Jacobsen, Mark R. 2013a. "Evaluating U.S. Fuel Economy Standards In a Model with Producer and Household Heterogeneity." *American Economic Journal: Economic Policy* 5 (2):148–187.
- . 2013b. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior." *American Economic Journal: Applied Economics* 5 (3).
- Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem. 2015. "The Implications of Heterogeneity for the Regulation of Energy-Consuming Durable Goods." Manuscript: University of Chicago.
- Jacobsen, Mark R. and Arthur A. van Benthem. 2013. "Vehicle Scrappage and Gasoline Policy." NBER Working Paper 19055.
- Jaffe, Adam B. and Robert N. Stavins. 1994. "The Energy-Efficiency Gap." *Energy Policy* 22 (10):804–810.
- Kaplow, Louis. 2013. "Optimal Regulation with Exemptions and Corrective Taxes." Working Paper: Harvard University.
- Kleven, Henrik J., Camille Landais, Emmanuel Saez, and Esben Schultz. 2014. "Migration and Wage Effects of Taxing Top Earners: Evidence from Foreigners' Tax Scheme in Denmark." *Quarterly Journal of Economics* 129 (1):333–378.
- 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.
- Kniesner, Thomas J. and John D. Leeth. 1991. "Compensating wage differentials for fatal injury risk in Australia, Japan, and the United States." *Journal of Risk and Uncertainty* 4 (1):75–90.
- Kopczuk, Wojciech. 2003. "A Note on Optimal Taxation in the Presence of Externalities." *Economics Letters* 80:81–86.
- Leard, Benjamin, Joshua Linn, and Virginia McConell. 2016. "Fuel Prices, New Vehicle Fuel Economy, and Implications for Attribute-Based Standards." Resources for the Future Discussion Paper 16-04.
- Leard, Benjamin and Virginia McConell. 2015. "New Markets for Pollution and Energy Efficiency: Credit Trading under Automobile Greenhouse Gas and Fuel Economy Standards." Resources for the Future Discussion Paper 15-16.
- Mendelsohn, Robert. 1986. "Regulating Heterogeneous Emissions." *Journal of Environmental Economics and Management* 13:301–312.
- Nelson, Randy A., Tom Tietenberg, and Michael R. Donihue. 1993. "Differential Environmental Regulation: Effects on Electric Utility Capital Turnover and Emissions." *Review of Economics and Statistics* 75 (2):368–373.
- Reynaert, Mathias. 2015. "Abatement Strategies and the Cost of Environmental Regulation: Emission Standards on the European Car Market." Manuscript: KU Leuven.

- Reynaert, Mathias and James M. Sallee. 2016. “Corrective Policy and Goodhart’s Law: The Case of Carbon Emissions from Automobiles.” Working Paper: University of California at Berkeley.
- 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.
- 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.
- Sallee, James M., Sarah E. West, and Wei Fan. 2016. “Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations.” *Journal of Public Economics* 135:61–73.
- Slemrod, Joel. 2010. “Buenas Notches: Lines and Notches in Tax System Design.” Manuscript: University of Michigan.
- Sneeringer, Stacy and Nigel Key. 2011. “Effects of Size-Based Environmental Regulations: Evidence of Regulatory Avoidance.” *American Journal of Agricultural Economics* 93 (4):1189–1211.
- Spence, Michael A. 1975. “Monopoly, Quality, and Regulation.” *The Bell Journal of Economics* 6 (2):417–429.
- Stavins, Robert N. 2006. “Vintage-Differentiated Environmental Regulation.” *Stanford Environmental Law Journal* 25 (29):29–63.
- Train, Kenneth. 2009. *Discrete choice methods with simulation*. Cambridge university press.
- 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.

# Online Appendices: Not for Publication

## Appendix A Proofs of propositions

**Proposition 1.** *Assume that there is competitive compliance trading. If welfare weights are uniform ( $\theta_n = 1 \forall n$ ), the optimal policy involves no attribute basing. The optimal attribute slope is:*

$$\sigma'(a_n)^* = 0 \forall a_n.$$

Under compliance trading, a single shadow price, denoted  $\lambda$  will prevail. Consumer  $n$ 's problem can be written:

$$\max_{a_n, e_n} U_n = F_n(a_n, e_n) + I_n - P(a_n, e_n) + \lambda \times (e_n - \sigma(a_n) - \kappa).$$

The first-order conditions are:

$$\frac{\partial U_n}{\partial a_n} = \frac{\partial F_n}{\partial a_n} - \frac{\partial P}{\partial a_n} - \lambda \sigma'(a_n) = 0 \quad (9)$$

$$\frac{\partial U_n}{\partial e_n} = \frac{\partial F_n}{\partial e_n} - \frac{\partial P}{\partial e_n} + \lambda = 0. \quad (10)$$

When  $\theta_n = 1 \forall n$ , the planner's direct allocation problem is:

$$\max_{a_n, e_n} W = \sum_{n=1}^N \{F_n(a_n, e_n) - C(a_n, e_n) + I_n\} + \phi \sum_{n=1}^N e_n.$$

The first-best optimization conditions are found by differentiation:

$$\frac{\partial W}{\partial a_n} = \frac{\partial F_n}{\partial a_n} - \frac{\partial C}{\partial a_n} = 0 \quad (11)$$

$$\frac{\partial W}{\partial e_n} = \frac{\partial F_n}{\partial e_n} - \frac{\partial C}{\partial e_n} + \phi = 0. \quad (12)$$

Under perfect competition, prices equal marginal costs. Then, it is apparent that the planner's and consumers' first-order conditions are identical if and only if  $\sigma'(a_n) = 0$  for all  $n$  and  $\lambda = \phi$ . The endogenous  $\lambda$  will be an increasing function of  $\kappa$  because of the convexity of the cost function. Thus, some value of  $\kappa$  exists for which  $\lambda = \phi$ . Choosing that value of  $\kappa$  and  $\sigma' = 0$  makes the consumers' first-order conditions identical to the planner's first-best. ■

**Corollary 1.** *Assume welfare weights are uniform ( $\theta_n = 1 \forall n$ ). The optimal subsidy involves no attribute basing. The optimal attribute slope is:*

$$\sigma'(a_n)^* = 0 \forall a_n.$$

The planner's conditions are the same as above. The consumer's problem is now:

$$\max_{a_n, e_n} U_n = F_n(a_n, e_n) + I_n - P(a_n, e_n) + s \times (e_n - \sigma(a_n)).$$

The first-order conditions are:

$$\frac{\partial U_n}{\partial a_n} = \frac{\partial F_n}{\partial a_n} - \frac{\partial P}{\partial a_n} - s\sigma'(a_n) = 0 \quad (13)$$

$$\frac{\partial U_n}{\partial e_n} = \frac{\partial F_n}{\partial e_n} - \frac{\partial P}{\partial e_n} + s = 0. \quad (14)$$

The consumer's conditions match the first-best if and only if  $\sigma'(a_n) = 0$  for all  $n$  and  $s = \phi$ . ■

**Proposition 2.** *Assume that there is no compliance trading. Then, even if welfare weights are uniform ( $\theta_n = 1 \forall n$ ), the optimal linear regulation generally involves attribute basing. If the constraint binds for all  $n$ , the optimal attribute slope satisfies:*

$$\hat{\sigma}^* = \frac{\text{cov}(\lambda_n, a_n)}{\phi \left( \frac{\partial \bar{a}}{\partial \hat{\sigma}} - \bar{a} \frac{\partial \bar{a}}{\partial \kappa} \right)},$$

which is not zero unless  $\lambda_n$  is uncorrelated with  $a_n$ .

The planner solves:

$$\max_{\hat{\sigma}, \kappa} W = \sum_{n=1}^N \{F_n(a_n, e_n) - C(a_n, e_n) + I_n\} + \phi \sum_{n=1}^N e_n.$$

The first-order condition with respect to  $\kappa$  is:

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

Using the optimality conditions from the consumer's problem, this can be rewritten as:

$$\sum_n \lambda_n \hat{\sigma} \frac{\partial a_n}{\partial \kappa} + (\phi - \lambda_n) \frac{\partial e_n}{\partial \kappa} = 0. \quad (15)$$

When the constraint is binding,  $e_n = \hat{\sigma} a_n + \kappa$ . Total differentiation of this constraint yields a rela-



relationship between  $\partial a_n/\partial \kappa$  and  $\partial e_n/\partial \kappa$ , namely that  $\partial e_n/\partial \kappa = \hat{\sigma} \cdot \partial a_n/\partial \kappa + 1$ . Using this substitution and rearranging equation 15 yields:

$$\bar{\lambda} = \phi \left( 1 + \underbrace{\hat{\sigma} \frac{\partial \bar{a}}{\partial \kappa}}_+ \right), \quad (16)$$

The first-order condition for  $\hat{\sigma}$  is:

$$\sum_n \left( \frac{\partial F_n}{\partial a_n} - \frac{\partial C}{\partial a_n} \right) \frac{\partial a_n}{\partial \hat{\sigma}} + \left( \frac{\partial F_n}{\partial e_n} - \frac{\partial C}{\partial e_n} + \phi \right) \frac{\partial e_n}{\partial \hat{\sigma}} = 0.$$

Substituting the consumer's optimality conditions yields:

$$\sum_n (\hat{\sigma} \lambda_n) \frac{\partial a_n}{\partial \hat{\sigma}} + (\phi - \lambda_n) \frac{\partial e_n}{\partial \hat{\sigma}} = 0.$$

Total differentiation of the constraint yields  $\partial e/\partial \hat{\sigma} = \hat{\sigma} \partial a/\partial \hat{\sigma} + a$ . Substitution yields:

$$\sum_n (\hat{\sigma} \lambda_n) \frac{\partial a_n}{\partial \hat{\sigma}} + (-\lambda_n + \phi) \left( \hat{\sigma} \frac{\partial a_n}{\partial \hat{\sigma}} + a_n \right) = 0.$$

Canceling terms yields:

$$\sum_n -\lambda_n a_n + \phi \hat{\sigma} \frac{\partial a_n}{\partial \hat{\sigma}} + \phi a_n = 0.$$

We then use the definition of the sample covariance of  $\lambda_n$  and  $a_n$  to rewrite  $\sum_n \lambda_n a_n$  as  $\sum_n (\lambda_n - \bar{\lambda})(a_n - \bar{a}) + n^{-1} \sum_n a_n \sum_n \lambda_n = n(\text{cov}(\lambda_n, a_n) + \bar{a}\bar{\lambda})$ , substitute equation 16 to rewrite the average shadow price, and rearrange. This yields the result.

It is apparent that  $\hat{\sigma}$  must be non-zero, unless the covariance between the shadow price and the attribute is zero under a flat standard. Otherwise, there is a contradiction. ■

**Corollary 2.** *Assume that there is no compliance trading, that welfare weights are uniform ( $\theta_n = 1 \forall n$ ), that the constraint binds for all  $n$ , and that there is a perfect correlation between attributes ( $e_n^0 = b + m a_n^0$  with  $m \neq 0$ ). With a uniform quadratic loss function for all  $n$ , the optimal linear regulation involves attribute basing but it does not fully equalize marginal costs, even though this is possible. The optimal attribute slope satisfies  $\hat{\sigma}^* \neq 0$  and  $\hat{\sigma}^* \neq m$ .*

The uniform quadratic loss function is important because it implies that marginal cost scales propor-

tionally to the distance between the privately optimal bundle and the standard. Any other loss function with that property will produce the result. With a uniform quadratic loss function, the shadow price is  $\xi(\hat{\sigma}a_n^0 + \kappa - e_n^0)$  (This is equation 4.) Substitute  $e_n^0 = b + ma_n^0$ . Then  $\lambda_n = \xi((\hat{\sigma} - m)a_n^0 + \kappa - b)$ .

Suppose  $\hat{\sigma} = m$ . Then,  $\lambda_n = \xi(\kappa - b)$  for all  $n$ . This demonstrates that it is possible to fully equalize marginal costs. When  $\hat{\sigma} = m$ , the covariance of  $a_n$  and  $\lambda_n$  (which does not vary) is zero. Plugging this into the result from Proposition 2 implies that  $m = 0$ , which is a contradiction.

For the other half of the result, suppose that  $\hat{\sigma} = 0$ . Then,  $cov(\lambda_n, a_n) = cov(\xi(-ma_n^0 + \kappa - b), a_n) = -m\xi cov(a_n^0, a_n)$ . Plugging this into the result from Proposition 2, and multiplying out the denominator against zero and dividing by  $-m\xi$  implies that  $0 = cov(a_n^0, a_n)$ , which is a contradiction because the starting and ending values of  $a_n$  are correlated.

To see the correlation, in the quadratic case, the consumer's optimization problem implies that the optimal change in  $a_n$  is  $\Delta a_n^* = \frac{2\beta\hat{\sigma} + \gamma}{2\beta\hat{\sigma}^2 + 2\gamma\hat{\sigma} + 2\alpha}(e_0 - \hat{\sigma}a_0 + \kappa)$ . When  $\hat{\sigma} = 0$  and  $e_n^0 = ma_n^0 + b$ , this reduces to  $\Delta a_n^* = \frac{m\gamma}{2\alpha}a_n^0 + \frac{\gamma}{2\alpha}(b - \kappa)$ . Thus,  $cov(a_n^0, a_n) = cov(a_n^0, a_n^0 + \Delta a_n) = cov(a_n^0, (m\gamma/(2\alpha) + 1)a_n^0) = (m\gamma/(2\alpha) + 1)var(a_n^0) \neq 0$  unless there is no variation in  $a_n^0$  (that is, all products are identical).■

**Proposition 3.** *Assume welfare weights are uniform ( $\theta_n = 1 \forall n$ ) and  $s = \phi$ . The deadweight loss from a subsidy with  $\sigma'(a_n) \neq 0$  is approximated as:*

$$DWL \approx \sum_n 1/2 \cdot \frac{\partial a_n}{\partial (s\sigma'(a_n))} (s\sigma'(a_n))^2.$$

Differentiating the planner's problem (equation 2) with respect to  $s\sigma'(a_n)$  yields the first-order condition:

$$\frac{\partial W}{\partial s\sigma'(a_n)} = \sum_n \left( \frac{\partial F_n}{\partial e_n} - \frac{\partial C}{\partial e_n} + \phi \right) \frac{\partial e_n}{\partial s\sigma'(a_n)} + \left( \frac{\partial F_n}{\partial a_n} - \frac{\partial C}{\partial a_n} \right) \frac{\partial a_n}{\partial s\sigma'(a_n)}$$

Substituting the consumer's optimality conditions yields, using  $s = \phi$ :

$$\frac{\partial W}{\partial s\sigma'(a_n)} = \sum_n (-\phi + \phi) \frac{\partial e_n}{\partial \phi\sigma'(a_n)} + (\phi\sigma'(a_n)) \frac{\partial a_n}{\partial \phi\sigma'(a_n)}.$$

The first term cancels.

Deadweight loss for a particular value of the slope  $\sigma'(a_n)$ , given  $s(= \phi)$ , is the integral of this derivative from 0 to  $s\sigma'(a_n) = \phi\sigma'(a_n)$ . We approximate the integral by assuming a constant derivative.

Integration then yields the result.

$$DWL = \sum_n \int_0^{\phi\sigma'(a_n)} \frac{\partial W}{\partial \phi\hat{\sigma}'(a)} d(\phi\hat{\sigma}'(a)) \approx \sum_n 1/2 \frac{\partial a_n}{\partial (\phi\sigma'(a_n))} (\phi\sigma'(a_n))^2.$$

We write the result using  $s$  in place of  $\phi$  to make the connection to tax wedges most clear. ■

**Proposition 4.** *Assume that welfare weights  $\theta_n$  vary. Then, the optimal linear subsidy involves attribute basing unless  $\theta_n$  is uncorrelated with  $a_n$ . The optimal attribute slope is:*

$$\hat{\sigma}^* = \left[ \left( \frac{\phi - s}{s} \right) \frac{\partial \bar{e}}{\partial \hat{\sigma}} - \text{cov}(\theta_n, a_n) \right] / \frac{\partial \bar{a}}{\partial \hat{\sigma}}.$$

Under lump-sum revenue recycling, each type will receive a tax equal to the opposite of the average subsidy. This tax is  $\tau = s(\bar{e}_n - \hat{\sigma}\bar{a}_n)$ . The planner's problem is thus:

$$\begin{aligned} \max_{\hat{\sigma}, s} W &= \sum_n \theta_n \{F_n(a_n, e_n) - C(a_n, e_n) + I_n - \tau + s(e_n - \hat{\sigma}a_n)\} + \phi \sum_n e_n \\ &= \sum_n \theta_n \{F_n(a_n, e_n) - C(a_n, e_n) + I_n + s(e_n - \bar{e}_n) - s\hat{\sigma}(a_n - \bar{a}_n)\} + \phi \sum_n e_n. \end{aligned}$$

Differentiating with respect to  $\hat{\sigma}$  and substituting in the consumer's optimality conditions yields this first-order condition:

$$\begin{aligned} \frac{\partial W}{\partial \hat{\sigma}} &= \sum_n \theta_n \left\{ -s \frac{\partial e_n}{\partial \hat{\sigma}} + s\hat{\sigma} \frac{\partial a_n}{\partial \hat{\sigma}} + s \left( \frac{\partial e_n}{\partial \hat{\sigma}} - \frac{\partial \bar{e}_n}{\partial \hat{\sigma}} \right) - s\hat{\sigma} \left( \frac{\partial a_n}{\partial \hat{\sigma}} - \frac{\partial \bar{a}_n}{\partial \hat{\sigma}} \right) - s(a_n - \bar{a}_n) \right\} + \phi \sum_n \frac{\partial e_n}{\partial \hat{\sigma}} \\ &= 0. \end{aligned}$$

The terms involving  $n$ -specific derivatives cancel. Dividing through by  $s$  and pulling the average derivative terms, which do not vary by  $n$ , out of the summation yields:

$$0 = -\frac{\partial \bar{e}}{\partial \hat{\sigma}} \sum_n \theta_n + \hat{\sigma} \frac{\partial \bar{a}}{\partial \hat{\sigma}} \sum_n \theta_n - \sum_n \theta_n (a_n - \bar{a}) + \frac{\phi}{s} \sum_n \frac{\partial e_n}{\partial \hat{\sigma}}$$

By construction, the mean of  $\theta$  is 1, so  $\sum_n \theta_n = N$ . Using that substitution, dividing through by  $N$ , and rewriting  $\sum_n \frac{\partial e_n}{\partial \hat{\sigma}}$  in terms of the mean derivative yields:

$$0 = \frac{s - \phi}{s} \frac{\partial \bar{e}}{\partial \hat{\sigma}} + \hat{\sigma} \frac{\partial \bar{a}}{\partial \hat{\sigma}} - \frac{1}{n} \sum_n \theta_n (a_n - \bar{a}).$$

Substitute using the definition of covariance (note that  $cov(\theta_n, a_n) = \sum_n (\theta_n - \bar{\theta})(a_n - \bar{a}) = \sum_n \theta_n (a_n - \bar{a})$ ). Solving for  $\hat{\sigma}$  yields the result. ■

## Appendix B Additional theoretical results

If a policy is constrained to include attribute basing, this influences the optimal stringency of policy because the wedge for  $a$  (and hence deadweight loss) is mechanically related to the wedge for  $e$ . Thus, if there is attribute basing, the second-best policy stringency will be attenuated away from the Pigouvian benchmark. This is demonstrated for the case of a linear subsidy in Proposition 5. Note that we use bars to denote sample averages of variables or derivatives (e.g.,  $N^{-1} \sum_n \partial e_n / \partial s \equiv \partial \bar{e} / \partial s$ ).

**Proposition 5.** *Assume that  $\hat{\sigma}$  is fixed and that welfare weights are uniform ( $\theta_n = 1 \forall n$ ). For a linear subsidy, the second-best subsidy rate is:*

$$s^* = \frac{\phi}{1 - \hat{\sigma} \left( \frac{\partial \bar{a} / \partial s}{\partial \bar{e} / \partial s} \right)} \leq \phi$$

The denominator of the expression for the optimal subsidy will be greater than one because the wedges in  $a$  and  $e$  will have the same sign when  $\hat{\sigma}$  is negative (and vice-versa).<sup>43</sup> Thus, the use of attribute basing in a policy implies that the second-best price on the externality-generating characteristic  $e$  is less than marginal benefits.

This attenuation grows as the market response to a policy tilts towards gaming the attribute and away from improving the targeted characteristic (that is, as  $\partial a / \partial s$  gets large relative to  $\partial e / \partial s$ ). In the limit, when actors respond to policy exclusively by manipulating the attribute, the optimal subsidy goes to zero. Thus, when the attribute responds more elastically to policy, deadweight loss will be larger (Proposition 3) and policy should be scaled back and made less stringent (Proposition 5). Related to these results, in our empirical analysis, we detect a large response in the attribute to policy and we estimate the proportion of compliance behavior that comes from changing the targeted characteristic versus the secondary attribute. Note that the optimal stringency  $\lambda$  of a regulation with trading will equal the same expression for  $s^*$  derived in the proposition.

---

<sup>43</sup>The wedge in  $e$  is  $s$  and the wedge in  $a$  is  $-s\hat{\sigma}$ .

**Proof:** Taking  $\hat{\sigma}$  as fixed, the planner's second-best choice of  $s$  solves:

$$\max_s W = \sum_{n=1}^N \{F_n(a_n, e_n) - C(a_n, e_n) + I_n\} + \phi \sum_{n=1}^N e_n.$$

The first-order condition is:

$$\sum_n \left( \frac{\partial F_n}{\partial e_n} - \frac{\partial C}{\partial e_n} + \phi \right) \frac{\partial e_n}{\partial s} + \left( \frac{\partial F_n}{\partial a_n} - \frac{\partial C}{\partial a_n} \right) \frac{\partial a_n}{\partial s} = 0.$$

Substituting the optimality conditions from the consumer's problem and pulling constants in front of the summation signs yields:

$$(-s + \phi) \sum_n \frac{\partial e_n}{\partial s} + (\hat{\sigma}s) \sum_n \frac{\partial a_n}{\partial s} = 0.$$

Use bars denote average derivatives (e.g.,  $N^{-1} \sum_n \partial e_n / \partial s \equiv \bar{\partial e} / \partial s$ ), and divide through by  $N$ . Rearranging yields the result.

The inequality follows because denominator is greater than 1 because the second term is negative. (We assume that  $\phi$  is positive. The sign is flipped if  $\phi$  is negative.) If  $\hat{\sigma}$  is negative, then the wedges in  $a$  and  $e$  will have opposite signs, so that the ratio of derivatives will be negative. If  $\hat{\sigma}$  is positive, then the wedges in  $a$  and  $e$  will have the same sign, so that the ratio of derivatives will be positive. ■

## B.1 Attribute-basing and second-best targeting

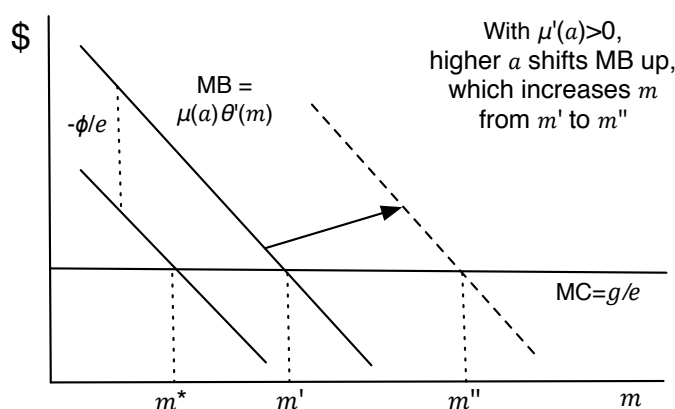
In our baseline model, the first-best Pigouvian tax is feasible. Might second-best considerations justify attribute-basing? In a second-best setting, we would not expect  $\sigma' = 0$ , in general. Distorting the choice of the attribute might help alleviate distortions on other margins, or attribute-basing might function as a tag in the spirit of Akerlof (1978). In practice, however, we believe that such considerations cannot justify existing attribute-based energy efficiency policies.

An energy-efficiency tax or subsidy scheme cannot truly be first-best because it will fail to correct the consumer's incentives regarding the intensity of use of the durable good. In fact, such policies tend to exacerbate the intensive use margin by lowering the cost of utilization. This is known in the literature as the rebound effect.<sup>44</sup> We modify our model to capture this additional margin for the case

---

<sup>44</sup>See Borenstein (Forthcoming) and Gillingham, Rapson, and Wagner (Forthcoming) for recent discussions.

**Figure A-1:** Graph of consumer's first-order condition for  $m$



The socially optimal  $m$  is achieved with a tax equal to  $\phi/e$ , which lowers  $m$  from  $m'$  to  $m^*$ . A policy that increases  $a$  will raise marginal benefits, raising  $m$  from  $m'$  to  $m''$ . Marginal cost may shift as well.

of a representative consumer. The consumer's maximization problem is:

$$\max_{a,e,m} U = \mu(a)\theta(m) - P(a,e) + I + se - s\sigma'a - \frac{gm}{e},$$

where  $m$  is intensity of use,  $\mu(a)\theta(m)$  is the utility derived from quality-adjusted usage,  $g$  is the cost of energy per unit, and  $\frac{gm}{e}$  is therefore the cost of usage. We assume that  $\mu' > 0$ ,  $\mu'' < 0$ ,  $\theta' > 0$  and  $\theta'' < 0$ . The assumption that  $\mu' > 0$  is key; it says that when  $a$  is higher, utilization is more valuable. This is the sign that we expect for energy-efficiency policies, where  $a$  represents some attribute of the product that makes it more desirable to use (e.g., conditional on cost, one will drive a larger car more because it is of higher quality). The planner's welfare function is  $W = \mu(a)\theta(m) - P(a,e) + I - \frac{gm}{e} - \frac{\phi m}{e}$ , where  $\phi$  is the externality per unit of energy used.

The consumer's first-order condition for utilization equates the cost of use with the benefit of use:  $\mu(a)\theta'(m) = g/e$ . This differs from the planner's condition, which is  $\mu(a)\theta'(m) = g/e + \phi/e$ , so that  $\phi/e$  represents the marginal externality from raising  $m$ , given a value of  $e$ . The consumer's choice problem is illustrated in Figure A-1. Conditional on  $e$ , the socially optimal  $m$  is below what the consumer will choose. The planner wishes to decrease  $m$ .

Can an attribute-based policy that raises  $a$  help alleviate this distortion? The first effect of inducing an increase in  $a$  is to shift the marginal benefit curve upward, which will unambiguously raise  $m$ . The second effect is that a change in  $a$  can induce a change in  $e$ , which would shift the marginal cost curve.

This effect can go in either direction. Thus, to lower  $m$ , a subsidy to  $a$  must induce a sufficiently large *fall* in efficiency  $e$  so that the cost of utilization rises. But, this is working against the goal of raising efficiency, and a lower  $e$  could be created by altering  $s$  without inducing a distortion to  $a$ . Instead, as long as  $\mu' > 0$ , the most likely outcome is that the second-best attribute slope is positive ( $\sigma'^{SB} > 0$ ); that is, there should be a tax on the attribute, instead of a subsidy, whenever the attribute is a quality that raises the value of using the good. We derive the second-best values of  $s$  and  $\sigma'$  in the appendix. The analysis shows that second-best targeting considerations can rationalize attribute-basing ( $\sigma' \neq 0$ ), but the empirically relevant cases imply that  $\sigma' > 0$ , which is the opposite of observed policy ( $\sigma' < 0$ ).

The representative consumer setup here abstracts from the possible benefits of the attribute as an Akerlof tag, but it is easy to see that actual attribute-based policies are poor tags. An attribute will be a useful tag to the extent that it is correlated with a product's externality, conditional on  $e$ . But, Jacobsen et al. (2015) show that vehicle characteristics like size have a very weak correlation with lifetime mileage. Moreover, actual policies have explicitly selected attributes that are tightly correlated with  $e$ , which limits the usefulness of the tag.<sup>45</sup>

As a result, we conclude that, while attribute-basing could play a role in second-best policies generally, second-best logic is unlikely to justify the real-world energy-efficiency regulations that we analyze in this paper.

Some additional algebraic detail is provided next. The first-order condition for  $s$  is:

$$\frac{\partial W}{\partial s} = (\mu'\theta - P'_a) \frac{\partial a}{\partial s} + \left(-P'_e + \frac{(g + \phi)m}{e^2}\right) \frac{\partial e}{\partial s} + \left(\mu\theta' - \frac{g + \phi}{e}\right) \frac{\partial m}{\partial s} = 0.$$

Substituting in the consumer's optimality conditions yields:

$$\frac{\partial W}{\partial s} = s\sigma' \frac{\partial a}{\partial s} + \left(-s + \frac{\phi m}{e^2}\right) \frac{\partial e}{\partial s} + \frac{-\phi}{e} \frac{\partial m}{\partial s} = 0. \quad (17)$$

Parallel steps yield the analogous first-order condition for  $\sigma'$ :

$$\frac{\partial W}{\partial \sigma'} = s\sigma' \frac{\partial a}{\partial \sigma'} + \left(-s + \frac{\phi m}{e^2}\right) \frac{\partial e}{\partial \sigma'} + \frac{-\phi}{e} \frac{\partial m}{\partial \sigma'} = 0. \quad (18)$$

Equations 17 and 18 have two unknowns ( $s$  and  $\sigma'$ ). Solving them involves rearrangement, substitution

---

<sup>45</sup>For example, U.S. regulators state that footprint's primary drawback is that it is less strongly correlated with fuel economy than is weight. See the Federal Register, volume 77, number 199, page 62687.

and simplification. Solving for the second-best subsidy rate ( $s^{SB}$ ) in those steps yields:

$$s^{SB} = \frac{m\phi}{e^2} - \frac{\phi}{e}d, \quad (19)$$

$$\text{where } d \equiv \left( \frac{\frac{\partial m/\partial\sigma'}{\partial a/\partial\sigma'} - \frac{\partial m/\partial s}{\partial a/\partial s}}{\frac{\partial e/\partial\sigma'}{\partial a/\partial\sigma'} - \frac{\partial e/\partial s}{\partial a/\partial s}} \right).$$

Equation 19 is the sum of two terms. The first is the marginal externality from increasing  $s$ ,  $m\phi/e^2$ , which is the first-best subsidy to  $s$  when  $m$  is fixed or there is a first-best tax on  $m$ . The second term is the marginal externality from  $m$  ( $\phi/e$ ) times a function of a collection of derivatives denoted  $d$ . This term captures the degree to which  $\sigma'$  versus  $s$  are effective tools for changing  $m$  versus  $a$  and  $e$ .

In turn, solving equation 18 for  $\sigma'$  yields:

$$\sigma'^{SB} = \left( 1 - \frac{m\phi}{e^2 s^{SB}} \right) \frac{\partial e/\partial\sigma'}{\partial a/\partial\sigma'} + \frac{\phi}{e s^{SB}} \frac{\partial m/\partial\sigma'}{\partial a/\partial\sigma'}. \quad (20)$$

Substituting in equation 19 yields a closed form result.

Even without the final substitution, we can see what is required to make  $\sigma'^{SB}$  negative, which is our primary concern. The second-best attribute slope has two terms. Each multiplies a factor that represents the distortion in one of the two margins involving the externality ( $s$  and  $m$ ) with a factor that is a ratio of derivatives indicating how much a change in  $\sigma'$  affects that variable versus  $a$ .

In our empirical results, we find estimates that imply that  $\partial e/\partial\sigma'$  is close to zero.<sup>46</sup> In that case, the first term of equation 20 will be close to zero, and the sign of the second term will determine the sign of  $\sigma'^{SB}$ .

As long as  $s$  is positive, then  $\phi/es > 0$ , so the sign of the second term depends solely on whether  $\partial m/\partial\sigma'$  and  $\partial a/\partial\sigma'$  have the same sign. If  $m$  and  $a$  are gross complements, so that increasing a subsidy to  $a$  increases the value of  $m$ , then this term will be positive. As argued in the main text, the direct effect of changing  $a$  on  $m$  is that it shifts the marginal benefits of utilization by, on the margin,  $\mu'(a)\theta(m)da$ . As long as  $\mu'(a) > 0$  (that is, products with more  $a$  are more desirable to use), this direct effect will be positive, which implies that  $\partial m/\partial\sigma'$  and  $\partial a/\partial\sigma'$  will have the same sign. Thus, the only way that  $m$  and  $a$  could be substitutes is if subsidizing  $a$  causes  $e$  to fall (thereby raising the cost of utilization on the margin) by a large enough amount to offset the direct effect of an increase in

---

<sup>46</sup>That is, subsidizing weight has a minimal net effect on fuel economy. This comes from our estimates showing that the interaction term on changes in  $e$  and  $a$  is approximately zero, which implies that the cross-partial in the utility and cost functions closely offset.



marginal benefits. (For the example of cars, a subsidy to weight would have to decrease fuel economy by a large enough amount to offset the increased marginal benefits of driving a larger car.)

As a result, the second term can be negative only when  $\partial e/\partial\sigma'$  is negative. But, our empirical estimates imply that this derivative is close to zero, which implies that  $\partial m/\partial\sigma' > 0$ . That is, in the empirically relevant case, the second-best policy will tax size. This is intuitive. The second-best policy will feature a subsidy to energy efficiency, and a tax on desirable attributes is used to mitigate the rebound effect.

## Appendix C Welfare implications of notched attribute-based policies

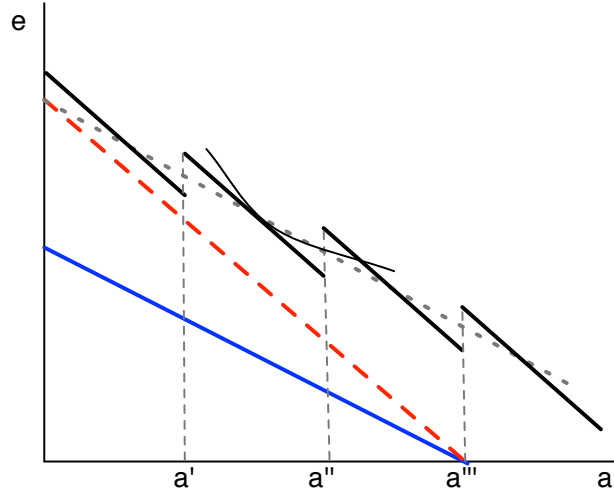
Our theory models smooth attribute-based functions, that is, cases where the target function  $\sigma(a_n)$  is everywhere differentiable in  $a$ . The Japanese policy that we analyze empirically has notches, so that  $\sigma(a_n)$  is a step-function. Here, we briefly argue that the welfare implications from our theory carry over to notched policies. We first consider “single notched” systems, like the Japanese corporate average fuel-economy standard, and then discuss “double notched” systems like the model-specific subsidy in Japan. In both cases, we discuss a subsidy policy rather than a regulation for notational ease.

How does a single notched policy, where  $\sigma(a_n)$  is a step function but the marginal incentive for  $e$  is smooth, affect choice? We provide initial intuition graphically. Figure A-2 shows an isocost curve, that is, the set of values of  $a$  and  $e$  for which a consumer spends a constant amount on the durable net of the subsidy,  $P(a_n, e_n) - s \times (e_n - \sigma(a_n))$ . 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  $\sigma(a_n)$  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 rep-

**Figure A-2:** Isocost Curve with a Notched Attribute-Based Subsidy



representing the Pigouvian subsidy (i.e, if  $S(a, e) = se$ ). 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 changed at all by the attribute basing 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  $a'$ ,  $a''$  or  $a'''$ . That is, all of the distortion is evident from those who “bunch” at the notch points.

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 subsidy jumps by amount  $\tau > 0$ . Then, the tax function can be written as:

$$S(a, e) = \begin{cases} s \cdot e & \text{if } a < a' \\ s \cdot e + \tau & \text{if } a \geq a'. \end{cases} \quad (21)$$

Denote by  $(a^*, e^*)$  the bundle chosen by a consumer facing a Pigouvian tax of  $s \cdot e$ . If the consumer’s choice under the smooth attribute policy had  $a^* > a'$ , then the addition of the notch  $\tau$  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  $\mathcal{L} = F(a, e) - P(a, e) + I - G + se + \tau + \mu[a - a']$ , 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,  $\mu$ , 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  $\tau$ ) 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:

$$\max_e = F(a', e) - P(a', e) + I - G + se + \tau, \quad (22)$$

which has the same first order condition for  $e$  as the case without  $\tau$ . Just as in the smooth case, any distortion to the choice of  $e$  comes through “general equilibrium” effects, through which a distortion in  $a$  shifts the marginal costs and benefits of  $e$ , which might result in a change in  $e$ .

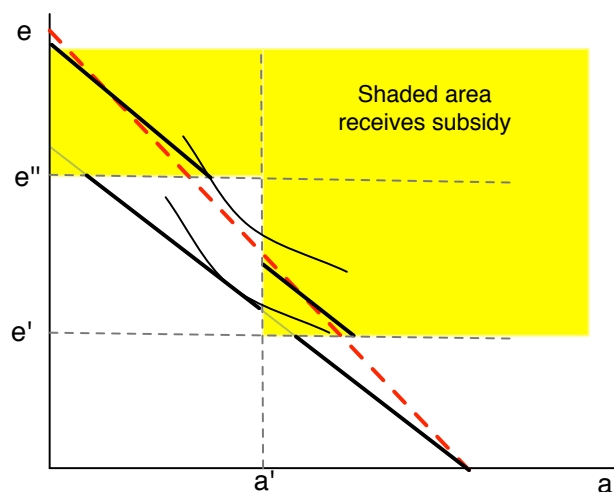
The distortion in  $a$  will be analogous to a traditional Harberger triangle and thus rising in  $\tau^2$ . The consumer will choose  $\tilde{a} = a'$  if and only if:

$$-\tau > P(a', \tilde{e}) - P(a^*, e^*) - (F(a', \tilde{e}) - F(a^*, e^*)), \quad (23)$$

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 has the same intuition as a traditional Harberger triangle, just as in the smooth case.

For our purposes, the 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.

**Figure A-3:** Isocost curve with notches for both  $a$  and  $e$



### C.1 Double notched policies

We next briefly describe the incentives created by a double notched policy, where the subsidy is not everywhere differentiable in  $e$  or  $a$ . 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 subsidy for such a system can be described algebraically as:

$$S(a, e) = \begin{cases} s_1 & \text{if } e > e'' \\ s_2 & \text{if } e'' > e > e' \text{ and } a > a' \\ 0 & \text{otherwise.} \end{cases} \quad (24)$$

An isocost curve for this case is shown in Figure A-3. 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  $s_1$  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  $s_1$  and  $s_2$  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  $s_1 \neq s_2$ . In many policy examples,  $s_1 = s_2$ ,

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 A-3, 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.

## Appendix D Panel Data Set for Section 4

We construct panel data of 675 cars by linking cars sold in 2008 (before the policy change) and 2012 (the last year of our data) based on a unique product identifier (ID) in the regulatory data. Product ID is narrower 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, each of which will have a unique ID.

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 exactly the same model name, displacement, drive type, and transmission. Entry and exit imply that not all data points are matched, and we end with 675 matched records. Our matching procedure guarantees that we match the same model name, which avoids mismatching the panel structure of the data. We take this approach because it provides transparent matching criteria.

A potential drawback of this approach is that firms may change some of their model names over time, yet they are targeting similar customer segments. To address this concern, we also conduct our analysis by including unmatched cars from the first matching criteria whenever we can match them using displacement, drive type, and transmission, while ignoring model names. This procedure produces a slightly different set of matched data, but our final estimation results are very similar regardless of which matching procedure is used.

## Appendix E Derivations of the Loss Functions in Section 4

This section provides a detailed description of how we derive the loss functions in Section 4. We analyze data before and after the policy change. For vehicle  $n$ , we denote  $a_n$  and  $e_n$  as the second-period characteristics and  $a_n^o$  and  $e_n^o$  as the first-period characteristics. We make the assumption of perfect competition. Then, using the notations in Section 2, the welfare for vehicle  $n$ , omitting regulatory incentives and dropping the numeraire, can be written as  $F_n(a_n, e_n) - C(a_n, e_n)$ .

First, consider a simple case, in which there is no fleet-average compliance regulation. Before the policy change, there is no regulation. After the policy change, there is a car-specific subsidy for cars that meet the standard ( $e_n \geq \sigma(a_n)$ ). Because the first-period characteristics  $a_n^o$  and  $e_n^o$  are at the private optimum, any deviation from that point creates a loss, which is  $L_n = F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)]$ . The second-period optimization problem for product  $n$  is then to choose the  $a_n$  and  $e_n$  values that maximize the loss plus the subsidy:

$$W_n = F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)] + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz},$$

where  $L_n = F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)] \leq 0$  and  $L_n$  is peaked at  $(a_n, e_n) = (a_n^o, e_n^o)$ . In another words, because the first-period characteristics  $a_n^o$  and  $e_n^o$  are at the private optimum, the loss function  $L_n$  would be the lowest and zero at  $(a_n^o, e_n^o)$ . This motivates us to begin with a quadratic functional form for  $L_n$  in our estimation. In the first specification in Section 4, we use a quadratic loss function:  $L_n = \alpha(a_n - a_n^o)^2 + \beta(e_n - e_n^o)^2 + \gamma(a_n - a_n^o)(e_n - e_n^o)$  and estimate:

$$W_n = \alpha(a_n - a_n^o)^2 + \beta(e_n - e_n^o)^2 + \gamma(a_n - a_n^o)(e_n - e_n^o) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz},$$

which is equation 7 in Section 4.

Second, consider the case with the presence of fleet-average compliance regulation. We denote  $\lambda$  and  $\lambda^o$  as the shadow prices of the fleet-average regulation at the second period and first period. Then, the payoff to first period choices is  $F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o) + \lambda^o \times (e_n^o - \sigma^o(a_n^o))$ . In the second period, there is a new fleet-average regulation and vehicle-specific subsidy policy. Then, the payoff to second-period choices is  $F_n(a_n, e_n) - C(a_n, e_n) + \lambda \times (e_n - \sigma(a_n)) + \tau \cdot 1(e_n \geq \sigma(a_n))$ . The second-period optimization problem for product  $n$  is then to choose the  $a_n$  and  $e_n$  values that maximize the objective

function:

$$\begin{aligned}
V_n &= F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)] \\
&\quad + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n^o - \sigma^o(a_n^o)) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz} \\
&= f_n + \lambda(e_n - \sigma(a_n)) - \lambda_n^o(e_n^o - \sigma^o(a_n^o)) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz},
\end{aligned}$$

where  $f_n \equiv F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)]$ . The problem with this equation is that  $f_n$  is not peaked at  $(a_n, e_n) = (a_n^o, e_n^o)$ .  $(a_n^o, e_n^o)$  is not the simple private optimum but rather the optimal bundle in the presence of the first-period fleet-average regulation. Therefore, it is problematic to use a quadratic function with a peak at  $(a_n^o, e_n^o)$  to approximate  $f_n$ .

We can address this problem by adding and subtracting  $\lambda^o(e_n - \sigma^o(a_n))$  from  $V_n$ . Note that this is a mixed object—it is the second-period choice of  $a$  and  $e$  put into the first-period policy function and shadow price:

$$\begin{aligned}
V_n &= F_n(a_n, e_n) - C(a_n, e_n) - [F_n(a_n^o, e_n^o) - C(a_n^o, e_n^o)] \\
&\quad + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n^o - \sigma^o(a_n^o)) + \tau \cdot 1(e_n \geq \sigma(a_n)) + \varepsilon_{nz} \\
&\quad + \lambda^o(e_n - \sigma^o(a_n)) - \lambda^o(e_n - \sigma^o(a_n)) \\
&= [F_n(a_n, e_n) - C(a_n, e_n) + \lambda^o(e_n - \sigma^o(a_n)) - F_n(a_n^o, e_n^o) + C(a_n^o, e_n^o) - \lambda^o(e_n^o - \sigma^o(a_n^o))] \\
&\quad + \tau \cdot 1(e_n \geq \sigma(a_n)) + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n - \sigma^o(a_n)) + \varepsilon_{nz} \\
&= g_n + \tau \cdot 1(e_n \geq \sigma(a_n)) + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n - \sigma^o(a_n)) + \varepsilon_{nz}
\end{aligned}$$

where  $g_n \equiv [F_n(a_n, e_n) - C(a_n, e_n) + \lambda^o(e_n - \sigma^o(a_n)) - F_n(a_n^o, e_n^o) + C(a_n^o, e_n^o) - \lambda^o(e_n^o - \sigma^o(a_n^o))]$ . Importantly,  $g_n$  is peaked at  $(a_n, e_n) = (a_n^o, e_n^o)$  because  $(a_n^o, e_n^o)$  is the optimum in the presence of the old policy. That is,  $g_n \leq 0$  is zero at  $(a_n^o, e_n^o)$ . That is, no changes in  $a$  and  $e$  would produce the lowest possible loss. This justifies us to have a quadratic functional form for  $g_n$  in our second specification in Section 4. We use a quadratic function:  $g_n = \alpha(a_n - a_n^o)^2 + \beta(e_n - e_n^o)^2 + \gamma(a_n - a_n^o)(e_n - e_n^o)$  and estimate:

$$\begin{aligned}
V_n &= \alpha(a_n - a_n^o)^2 + \beta(e_n - e_n^o)^2 + \gamma(a_n - a_n^o)(e_n - e_n^o) \\
&\quad + \tau \cdot 1(e_n \geq \sigma(a_n)) + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n - \sigma^o(a_n)) + \varepsilon_{nz},
\end{aligned}$$

which is equation (8) in Section 4.

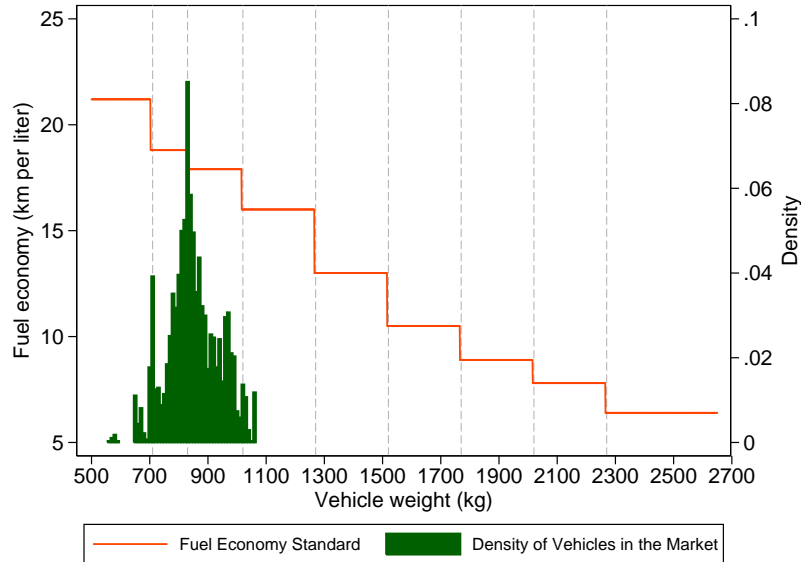
Finally, we can recover  $f_n$  from  $g_n$ . Note that  $g_n = f_n - [\lambda^o(e_n^o - \sigma^o(a_n^o)) - \lambda^o(e_n - \sigma^o(a_n))]$ . Therefore, once we have parameter estimates for  $g_n$  and  $\lambda^o$ , we can recover  $f_n = g_n + [\lambda^o(e_n^o - \sigma^o(a_n^o)) - \lambda^o(e_n - \sigma^o(a_n))]$ . This  $f_n$  provides the loss function for  $a_n$  and  $e_n$  excluding the effects of the old regulation.



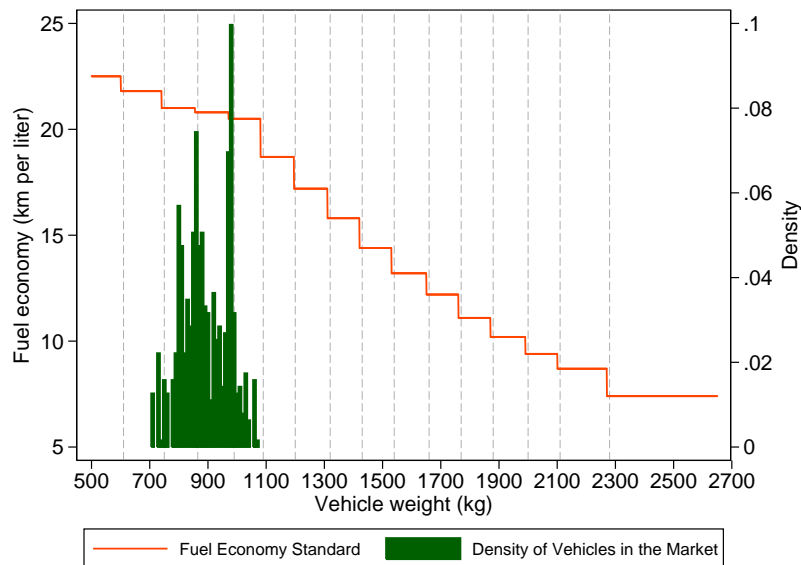
## Appendix F Additional Figures and Tables

**Figure A-4:** Fuel-Economy Standard and Histogram of Vehicles: Kei-Cars (small cars)

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



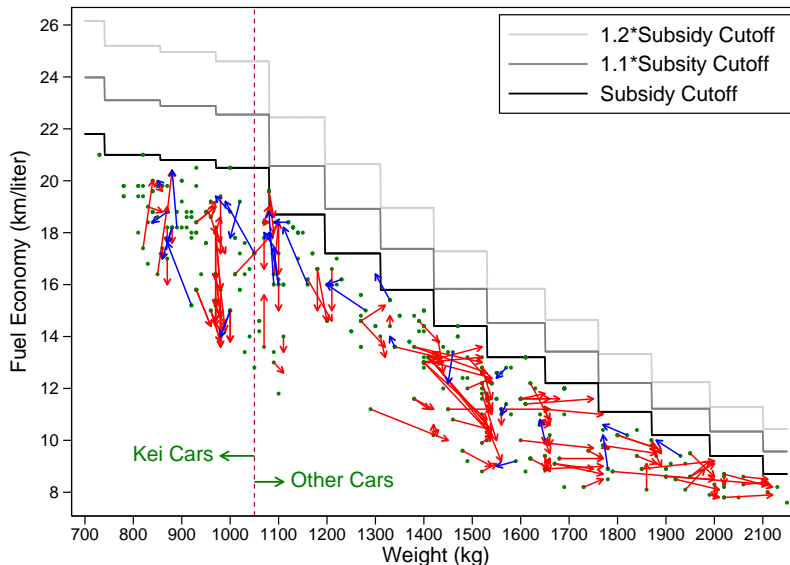
Panel B. Years 2009 to 2013 (New Fuel-Economy Standard Schedule)



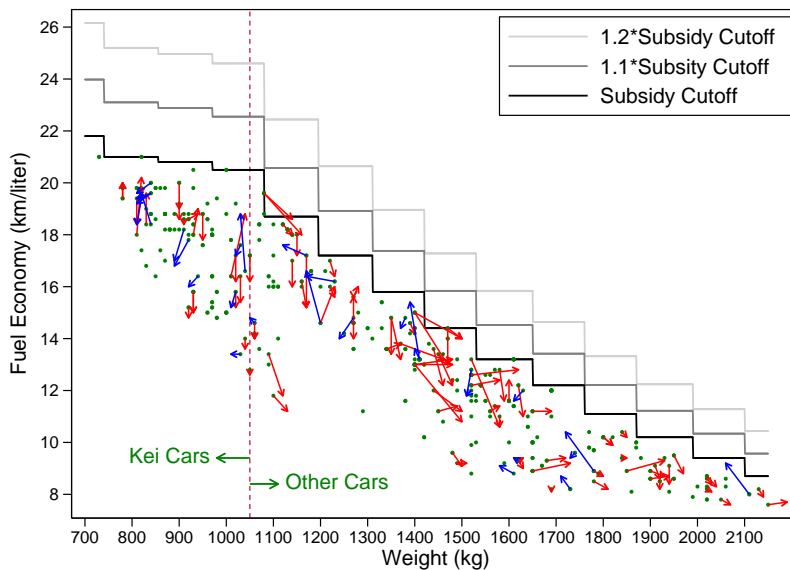
Note: “Kei-car” is a Japanese category of small vehicles; the displacement of kei-cars have to be less than 660 cc. Most kei-cars are not exported to other countries. Panel A shows the histogram of vehicles from 2001 to 2008, where all vehicles had the old fuel-economy standard. Panel B shows the histogram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

**Figure A-5:** Fuel Economy and Weight before and after the Policy Change for Vehicles that Did Not Receive a Subsidy

Panel A. Vehicles that did not receive a subsidy but bunched at weight notches



Panel B. Vehicles that did not receive a subsidy and did not bunch at weight notches



Note: This figure shows each vehicle’s fuel economy and weight before and after the introduction of the new subsidy that was applied to each vehicle individually. The scatterplot shows each car’s starting values of fuel economy and weight in 2008—the year before the policy change. We also show “arrows” connecting each car’s starting values in 2008 with its values in 2012. The figure also includes three step functions that correspond to the three tiers of the new incentive’s eligibility cutoffs.

**Table A-1:** Estimates of the Loss Function (Excluding Kei-cars)

	Logit (1)	Random-coefficient logit (2)
$\alpha : (\Delta \text{Weight})^2$	-0.71 (0.07)	-1.10 (0.15)
$\beta : (\Delta \text{Fuel consumption})^2$	-1.00 (0.08)	-1.30 (0.13)
$\gamma : \Delta \text{Weight} \times \Delta \text{Fuel consumption}$	0.19 (0.11)	0.26 (0.11)
$\tau: 1\{\text{Subsidy}\}$	0.48 (0.17)	0.62 (0.17)
Stadard deviation of random-coefficient		
$\alpha : (\Delta \text{Weight})^2$		0.42 (0.05)
$\beta : (\Delta \text{Fuel consumption})^2$		0.39 (0.05)
$\gamma : \Delta \text{Weight} \times \Delta \text{Fuel consumption}$		0.75 (0.12)

Note: As a robustness check for the results in Table 4, we estimate the logit and random-coefficients logit models by excluding kei-cars. There are small changes in the coefficients, but it does not significantly change our main findings.

**Table A-2:** Counterfactual Policy Simulations (Excluding Kei-cars)

	$\Delta e$ : Fuel consumption (liter/100km)	$\Delta a$ : Weight (kg)	$\Delta L$ Welfare loss (\$/car)	$\Delta L$ relative to efficient	S.D. of MC (\$/car)	Additional loss from safety externality (\$/car)
Panel A: Based on logit estimates						
Efficient	-0.71	-9.50	893	1.00	0	-175
Flat	-0.71	-12.81	3613	4.05	4397	-236
ABR	-0.71	39.42	2277	2.55	2215	726
Panel B: Based on random-coefficient logit estimates						
Efficient	-0.71	-10.48	821	1.00	0	-193
Flat	-0.71	-10.04	3477	4.23	4311	-185
ABR	-0.71	45.94	2460	3.00	2417	846

Note: As a robustness check for the results in Table 6, we provide the same policy simulations by excluding kei-cars. There are small changes in the coefficients, but it does not significantly change our main findings.

**Table A-3:** Counterfactual Policy Simulations (With Firm Dummy Variables Interacted with Cost Variables)

	$\Delta e$ : Fuel consumption (liter/100km)	$\Delta a$ : Weight (kg)	$\Delta L$ Welfare loss (\$/car)	$\Delta L$ relative to efficient	S.D. of MC (\$/car)	Additional loss from safety externality (\$/car)
Efficient	-0.79	-9.29	951	1.00	0	-171
Flat	-0.79	-40.94	6894	7.25	4457	-754
ABR	-0.79	26.07	2280	2.40	2467	480

Note: As a robustness check for the results in Table 6, we provide the same policy simulations based on a random-coefficients model in which all cost variables are interacted with firm dummy variables. There are small changes in the coefficients, but it does not significantly change our main findings.

**Table A-4:** Percent of Bunched Observations: Actual v.s. Predicted by the Loss Function (Logit)

	All	860 kg	980 kg	1090 kg	1200 kg	1320 kg	1430 kg
Actual (%)	27.60	1.58	3.17	1.81	0.90	1.58	1.81
Predicted (%)	32.73	1.82	2.50	4.32	2.27	2.50	5.68
	1540 kg	1660 kg	1770 kg	1880 kg	2000 kg	2110 kg	2280 kg
Actual (%)	4.52	5.66	2.71	1.36	1.36	1.13	0.00
Predicted (%)	7.27	0.91	0.91	2.27	1.14	0.68	0.00

Note: We use the estimated loss function in Table 4 and data from 2008 to predict bunching in 2012. Using the functional form assumption for  $L(\Delta a_n, \Delta e_n)$ , the estimated coefficients, and data from 2008 ( $a_n^o, e_n^o$ ), we calculate the predicted value for the value function  $V_{nz}$  and let each  $n$  to choose its optimal pair of  $(a_n, e_n)$  in 2012. We then compare the predicted bunching from this procedure to the actual bunching in the 2012 data.