Online Appendices: Not for Publication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 λ 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:

$$\frac{\partial U_n}{\partial a_n} = \frac{\partial F_n}{\partial a_n} - \frac{\partial P}{\partial a_n} - \lambda \sigma'(a_n) = 0 \tag{8}$$

$$\frac{\partial U_n}{\partial e_n} = \frac{\partial F_n}{\partial e_n} - \frac{\partial P}{\partial e_n} + \lambda = 0. \tag{9}$$

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:

$$\frac{\partial W}{\partial a_n} = \frac{\partial F_n}{\partial a_n} - \frac{\partial C}{\partial a_n} = 0 \tag{10}$$

$$\frac{\partial W}{\partial e_n} = \frac{\partial F_n}{\partial e_n} - \frac{\partial C}{\partial e_n} + \phi = 0. \tag{11}$$

Under perfect competition, prices equal marginal costs. Then, it is apparent that the plan-

ner's and consumers' first-order conditions are identical if and only if $\sigma'(a_n) = 0$ for all n and $\lambda = \phi$. The endogenous λ will be an increasing function of κ because of the convexity of the cost function. Thus, some value of κ exists for which $\lambda = \phi$. Choosing that value of κ and $\sigma' = 0$ makes the consumers' first-order conditions identical to the planner's first-best. ■

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:*

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

which is not zero unless λ_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 κ 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:

$$\sum_n \lambda_n \hat{\sigma}\frac{\partial a_n}{\partial \kappa} + (\phi - \lambda_n)\frac{\partial e_n}{\partial \kappa} = 0. \tag{12}$$

When the constraint is binding, $e_n = \hat{\sigma} a_n + \kappa$. Total differentiation of this constraint yields a relationship between $\partial a_n/\partial \kappa$ and $\partial e_n/\partial \kappa$, namely that $\partial e_n/\partial \kappa = \hat{\sigma} \cdot \partial a_n/\partial \kappa + 1$. Using

this substitution and rearranging equation 12 yields:

$$\bar{\lambda} = \phi \left(1 + \underbrace{\hat{\sigma} \frac{\partial \bar{a}}{\partial \kappa}}_{+} \right), \tag{13}$$

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 λ_n and a_n to rewrite $\sum_n \lambda_n a_n$ as $\sum_n (\lambda_n - \bar{\lambda})(a_n - \bar{a}) + n^{-1} \sum_n a_n \sum_n \lambda_n = n(cov(\lambda_n, a_n) + \bar{a}\bar{\lambda})$, substitute equation 13 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 1. *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 proportionally 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 18).) 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 λ_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 ϕ to make the connection to tax wedges most clear.

■

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

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

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:

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

Differentiating with respect to $\hat{\sigma}$ and substituting in the consumer's optimality conditions

yields this first-order condition:

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

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 θ 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

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 2 restates Proposition 1 for this case.

Corollary 2. *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 (**?**). Our result is a manifestation of the additivity property, and will thus hold in a wide range of second-best settings.

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).[8] 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

[8] The wedge in e is s and the wedge in a is $-s\hat{\sigma}$.

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 λ of a regulation with trading will equal the same expression for s^* derived in the proposition.

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., $N^{-1} \sum_n \partial e_n / \partial s \equiv \partial \bar{e} / \partial s$), 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 ϕ is positive. The sign is flipped if ϕ 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 **?**. 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.[9] We modify our model to capture this additional margin for the case of a representative consumer. The consumer's maximization problem is:

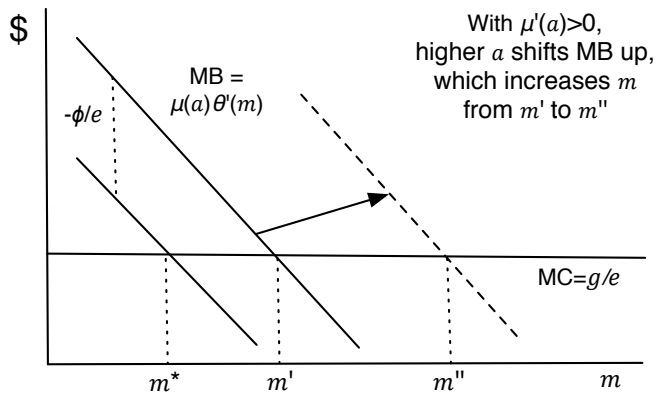
$$\max_{a,e,m} U = \mu(a)\theta(m) - P(a,e) + I + se - s\sigma' a - \frac{gm}{e},$$

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 ϕ 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 $\mu(a)\theta'(m) =$

[9] See **?** and **?** 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 ϕ/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.

$g/e + \phi/e$, so that ϕ/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 σ' 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, **?** 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.[10]

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:

$$\frac{\partial W}{\partial s} = s\sigma'\frac{\partial a}{\partial s} + \left(-s + \frac{\phi m}{e^2}\right)\frac{\partial e}{\partial s} + \frac{-\phi}{e}\frac{\partial m}{\partial s} = 0. \tag{14}$$

Parallel steps yield the analogous first-order condition for σ':

$$\frac{\partial W}{\partial \sigma'} = s\sigma'\frac{\partial a}{\partial \sigma'} + \left(-s + \frac{\phi m}{e^2}\right)\frac{\partial e}{\partial \sigma'} + \frac{-\phi}{e}\frac{\partial m}{\partial \sigma'} = 0. \tag{15}$$

Equations 14 and 15 have two unknowns (s and σ'). Solving them involves rearrangement, substitution and simplification. Solving for the second-best subsidy rate (s^{SB}) in those steps

[10]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.

yields:

$$s^{SB} = \frac{m\phi}{e^2} - \frac{\phi}{e}d, \tag{16}$$
$$\text{where } d \equiv \left(\frac{\frac{\partial m/\partial \sigma'}{\partial a/\partial \sigma'} - \frac{\partial m/\partial s}{\partial a/\partial s}}{\frac{\partial e/\partial \sigma'}{\partial a/\partial \sigma'} - \frac{\partial e/\partial s}{\partial a/\partial s}} \right).$$

Equation 16 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 (ϕ/e) times a function of a collection of derivatives denoted d. This term captures the degree to which σ' versus s are effective tools for changing m versus a and e.

In turn, solving equation 15 for σ' yields:

$$\sigma'^{SB} = \left(1 - \frac{m\phi}{e^2 s^{SB}}\right) \frac{\partial e/\partial \sigma'}{\partial a/\partial \sigma'} + \frac{\phi}{e s^{SB}} \frac{\partial m/\partial \sigma'}{\partial a/\partial \sigma'}. \tag{17}$$

Substituting in equation 16 yields a closed form result.

Even without the final substitution, we can see what is required to make σ'^{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 σ' affects that variable versus a.

In our empirical results, we find estimates that imply that $\partial e/\partial \sigma'$ is close to zero.[11] In that case, the first term of equation 17 will be close to zero, and the sign of the second term will determine the sign of σ'^{SB}.

As long as s is positive, then $\phi/es > 0$, so the sign of the second term depends solely on

[11] 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.

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 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 Additional commentary (removed for space)

A considerable amount of material was cut in order to meet the journal's page limits for print publication. This includes theoretical results, empirical results, references to the literature and commentary on our analysis. Subsequent appendices include additional figures and tables. Here we include the forms of commentary that we were forced to omit from the main text.

C.1 Attribute-based policies in the real world

There are many examples of attribute-based regulation that we did not have space to discuss. 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. In terms of consumer-facing labels, 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".

C.2 Literature

There is a literature on regulations that exempt firms based on size. Methodologically, our paper is quite distinct from this literature, but it is relevant to mention as it pertains to studies of an attribute-based regulation. Empirical results from this literature have been mixed in their search for distortion in firm size (Becker and Henderson 2000; **?**; **?**). The theoretical literature has focused on models of imperfect targeting of the externality (**??**) or per-firm administrative costs (**?**). 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 studying notched policies aimed at externalities, our work also relates to the literature on notched corrective taxation, which began with **?**, includes prior analysis of automobile fuel economy in **?**, 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.

In the main text, we mention only the most closely related studies that use bunching methods, but this is a rapidly growing literature, with significant contributions made also by **???** and **?**.

C.3 Comments on our model

In the model we introduce the constant κ implies that a linear policy can correct average prices, which distinguishes our policy from some performance standards. 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 additional appendix below. In the main text, we present the smooth policy version believing it to be more intuitive and general.

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. We explicitly model the two extremes of product-specific rules and market-wide averaging. Fleet averaging will achieve some of the marginal cost equilization. The more it does, the weaker is the efficiency case for attribute-basing.

Where we model trading, one might be concerned about efficiency in such a market with a small number of firms, but **?** discuss evidence that trades have taken place in the early years of CAFE trading.

C.4 Caveats to our model

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.[12]

[12]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

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.[13] 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 tax or registration fee that shifts all prices equally.

Third, advocates of attribute basing in car markets have argued that it promotes technology adoption.[14] 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.

---

on the logic of **?**, we conjecture that the key issue is whether the marginal consumer for each product values the attribute more or less than average.

[13]**?** treat this issue in detail for the case of a low-carbon fuel standard. **?** suggest that, for the case of automobiles, the benefits resulting from higher new car prices will be partly offset by changes in vehicle scrappage.

[14]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 (2013) for a related model that concludes that footprint-based CAFE is roughly safety neutral.

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.[15]

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 ω upon which policy cannot be based. Then, it might be useful to base policy on a secondary attribute a. **?** 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 they are correlated with income. The secondary attribute can be a tag in this setting. In this case, the welfare improvement from tagging will

[15]We are especially grateful to Ryan Kellogg for suggesting this line of reasoning. Consistent with this, **?** suggest that fleet fuel economy appears to be less responsive to gasoline price fluctuations since the introduction of attribute-based standards.

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.[16] 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.[17]

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 **?**). 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.[18]

[16] 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.

[17] **?** 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.

[18] See **?** for a review and **???** and **?** for evidence.

C.5 Comments on our graphical representation of theory

In the graphical subsection of our theory, we discuss the quadratic example. A quadratic loss function implies that the shadow price of a policy for type n can be written in closed form as:

$$\lambda_n = \xi \times (\hat{\sigma} a_n^0 + \kappa - e_n^0) \tag{18}$$
$$\text{where } \xi = \frac{4\alpha\beta - \gamma^2}{\beta\hat{\sigma}^2 + \gamma\hat{\sigma} + \alpha}.$$

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 ($e_n^0 = a + m a_n^0$) and $\hat{\sigma} = m$, marginal cost is equal to $\xi \times (\kappa - b)$ for all n. This leads to our graphical depictions.

C.6 Comments on Japanese policy

Note that we do find that some of 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 in the panel section controls for the effects of the corporate average fuel-economy standards.

Technically, the fuel-economy obligation extends to each weight segment separately in Japan. 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 the older version of the U.S. CAFE program, where there is one firm-wide requirement (but no trading across firms).

Also, the fuel-economy regulations are technically only binding in particular compliance years. This is different from the U.S. CAFE program, which requires compliance annually. 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.

The weight notches in the Japanese regulations figure prominently in our empirical strategy. One might be concerned about some endogenous choice of where to put the thresholds. 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.

The Japanese system operates through the front-runner system mentioned in the main text. 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.

C.7 Comments on data

We treat each type of vehicle as one observation in our analysis. 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.

We omit kei-cars from most of our main results because we believe this is functionally a different product market. 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 market segment, have different tax and insurance regulations, and are generally viewed as a distinct product category by Japanese consumers.

In figures we show 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.

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.

We use fuel consumption (l/100km) rather than fuel economy (km/liter) in our panel analysis 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).

We construct panel data of 439 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 439 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.

C.8 Comments on bunching analysis

The regulation we study 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.

In defining weight, 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.

We discuss the counterfactual interpretation of our bunching estimates. To elaborate, 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.

In our main estimates, we assume that weight manipulation is limited to a single step in the regulation, which is a conservative assumption. 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.

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.

In discussing safety externalities, we introduce a second externality. Note that 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.

Our safety analysis assumes that the accident risk is unpriced by the market. In principle, the externality risk may be partly priced through insurance or legal liability. **?**, however, argues that neither tort liability nor mandatory liability insurance prices safety externalities. In brief, tort liability requires negligence, not just that one be driving a dangerous vehicle. Liability insurance generally coves the cost of damages to a vehicle, but it is but a small

fraction of the value of a life. In addition, rate differences across vehicles are very coarse and reflect average driver characteristics in concert with vehicle attributes.

C.9 Comments on our panel analysis

The interpretation of our loss function is especially clear when consumer types are each matched to a specific vehicle. 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 **?** for a closely related model).

We do not estimate changes in markups in our empirical analysis. 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 **?**, 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.

Though it is not our focus, our discrete choice model estimates do deliver 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 (**?**) or uses static structural models (**???**). The estimate in column 3 implies a shadow price of \$1,162 ($= .45/.37 * \$1000$) per unit of l/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 **?** and **?** for the United Sates.

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.

In the case of a quadratic loss function that we analyze, 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.

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.

C.10 Validation of our discrete choice model estimates

In this subsection, 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 3 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.

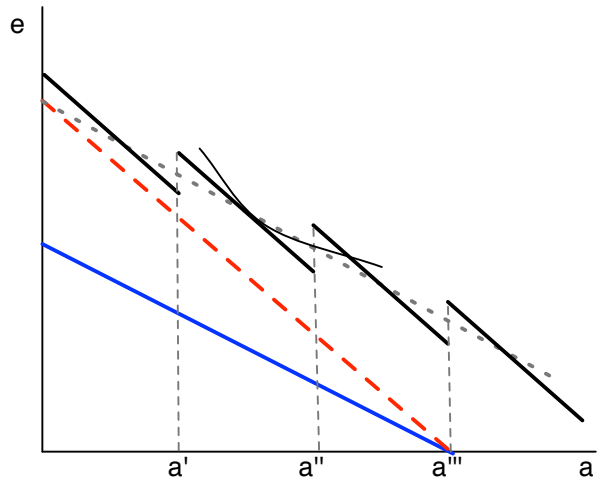
Results are in Table A.5 for the random-coefficients model and Appendix Table A.6 for the logit model. In the data, we observe that 27.6% of cars in 2012 are located at one of the weight notches in the fuel economy standard. With the logit specification, our model predict that 32.73% of cars bunch at 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.

Appendix D 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

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



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

$$S(a, e) = \begin{cases} s \cdot e & \text{if } a < a' \\ s \cdot e + \tau & \text{if } a \geq a'. \end{cases} \tag{19}$$

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 τ is purely an income effect. It has not changed the marginal price of a or e relative to each other or relative to x. Given quasi-linearity, this means that the durable choice of a consumer with $a^* > a'$ is unaffected by the introduction of a notched attribute policy.

When $a^* < a'$, the consumer will face a discrete choice of maintaining their original allocation or switching to a' exactly. They will not choose $a > a'$. To see why, suppose that they chose a value under the notched policy, call it $\tilde{a}$ strictly greater than a'. Then their optimization problem can be written $\mathcal{L} = F(a, e) - P(a, e) + I - G + se + \tau + \mu[a - a']$, 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, μ, is zero. In that case, the first-order conditions of the problem will be exactly the same as in the benchmark case with no attribute notch, which by construction featured an optimal choice of $a^* < a'$.

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

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

$$\max_e = F(a', e) - P(a', e) + I - G + se + \tau, \tag{20}$$

which has the same first order condition for e as the case without τ. 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 τ^2. The consumer will choose $\tilde{a} = a'$ if and only if:

$$-\tau > P(a', \tilde{e}) - P(a^*, e^*) - (F(a', \tilde{e}) - F(a^*, e^*)), \tag{21}$$

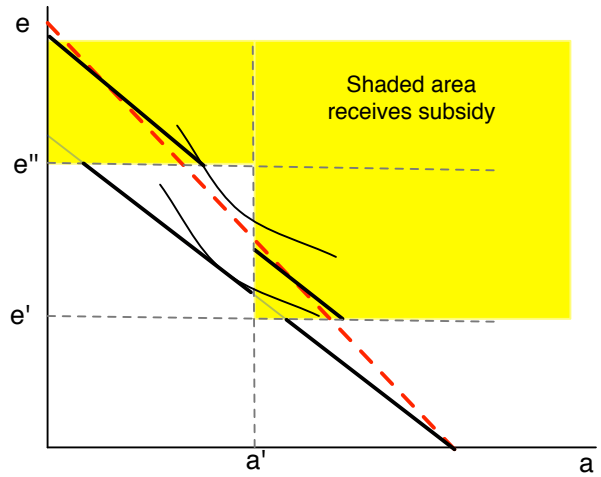
that is, whenever the tax benefit is larger than the cost increase from moving from (a^*, e^*) to $(a', \tilde{e})$ minus the increase in utility from that change. The welfare loss can be written as a Taylor expansion, which has the same intuition as a traditional Harberger triangle, just as in the smooth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

D.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

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



such a system can be described algebraically as:

$$S(a,e) = \begin{cases} s_1 & \text{if } e > e'' \\ s_2 & \text{if } e'' > e > e' \text{ and } a > a' \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

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 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 notation 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 6 in Section 4.

Second, consider the case with the presence of fleet-average compliance regulation. We denote λ and λ^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:

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

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:

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

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:

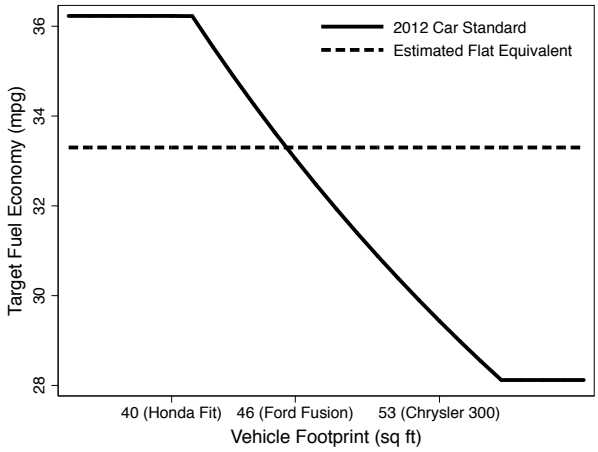
$$
\begin{aligned}
V_n =& \alpha(a_n - a_n^o)^2 + \beta(e_n - e_n^o)^2 + \gamma(a_n - a_n^o)(e_n - e_n^o) \\
&+ \tau \cdot 1(e_n \geq \sigma(a_n)) + \lambda(e_n - \sigma(a_n)) - \lambda^o(e_n - \sigma^o(a_n)) + \varepsilon_{nz},
\end{aligned}
$$

which is equation (7) 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 λ^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

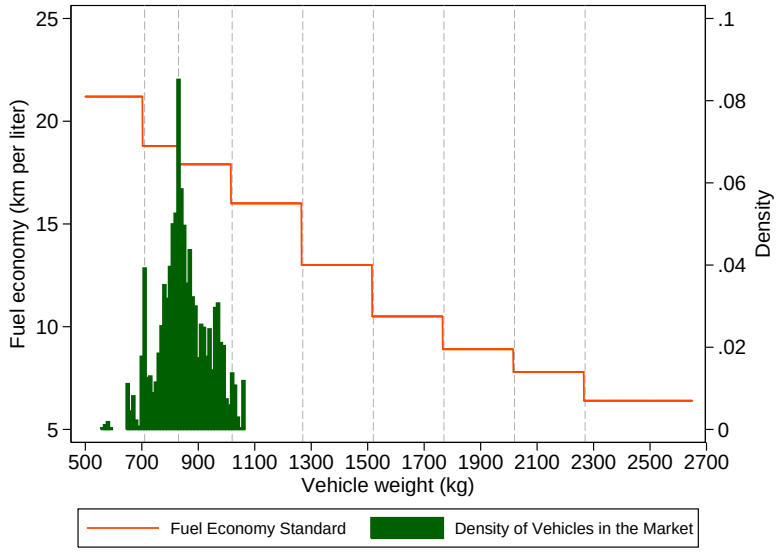
Figure A.1: Example of an Attribute-Based Regulation: U.S. CAFE standard 2012



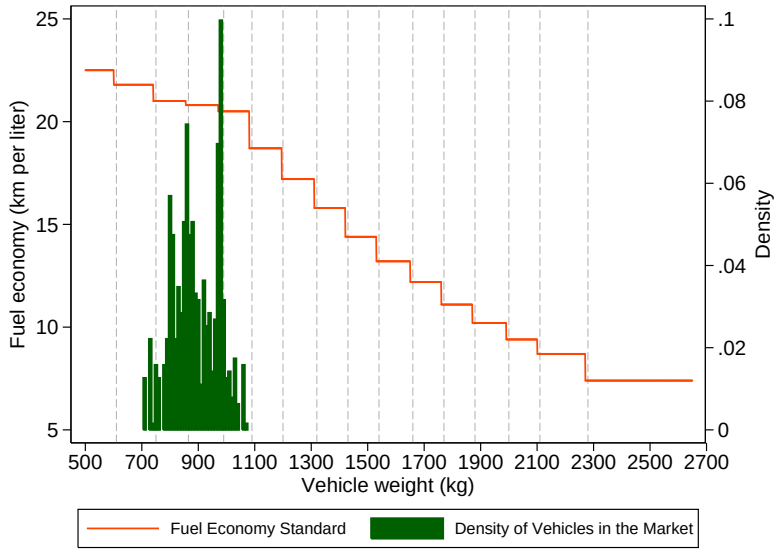
Solid line depicts the target function, which is a downward-sloping function of vehicle size, as measured by footprint (the square area trapped by the vehicle's tires). Firms must comply with the target function on average across their fleet. Example vehicles labeled on x-axis.

Figure A.2: Fuel-Economy Standard and Histogram of Vehicles: Kei-Cars (small cars)

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



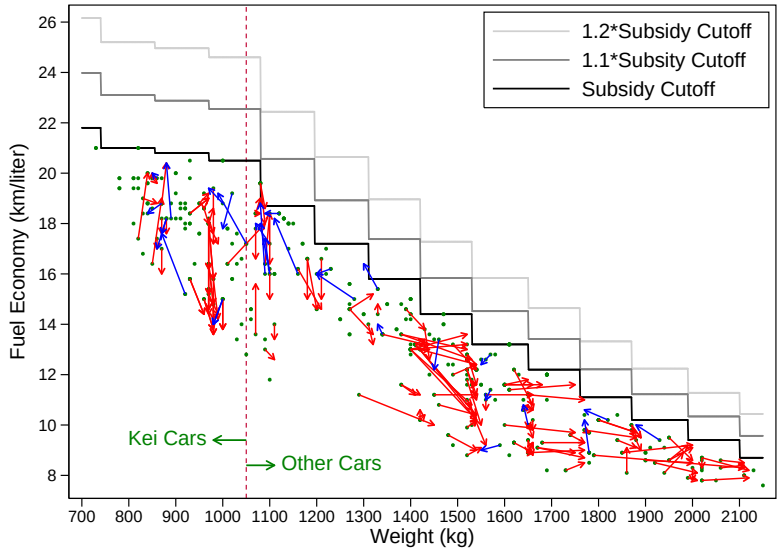
Panel B. Years 2009 to 2013 (New Fuel-Economy Standard Schedule)



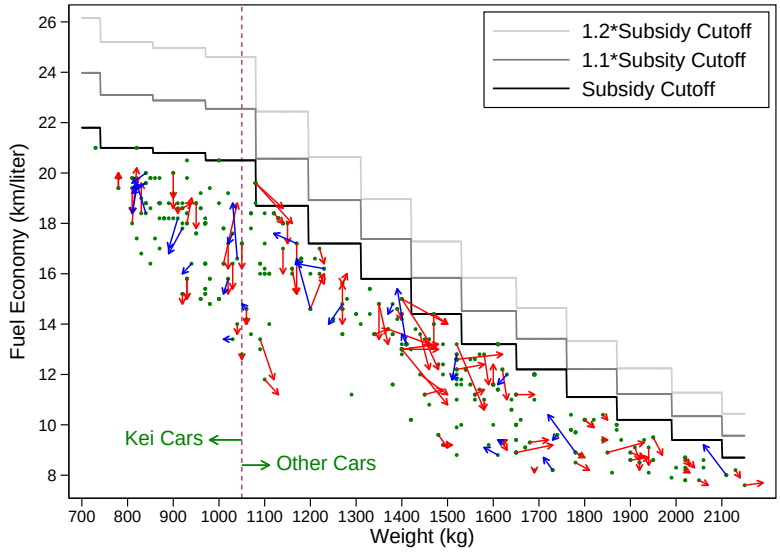
Note: "Kei-car" is is a Japanese category of small vehicles; the displacement of kei-cars have to be less than 660 cc. Most kei-cars are not exported to other countries. Panel A shows the histogram of vehicles from 2001 to 2008, where all vehicles had the old fuel-economy standard. Panel B shows the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

Figure A.3: 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: 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.

Table A.2: Estimates of the Loss Function (Excluding Kei-cars)

	Logit (1)	Random-coefficient logit (2)
$\alpha : (\Delta\text{Weight})^2$	-0.71	-1.10
	(0.07)	(0.15)
$\beta : (\Delta\text{Fuel consumption})^2$	-1.00	-1.30
	(0.08)	(0.13)
$\gamma : \Delta\text{Weight} \times \Delta\text{Fuel consumption}$	0.19	0.26
	(0.11)	(0.11)
τ: 1{Subsidy}	0.48	0.62
	(0.17)	(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 3, 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.3: Counterfactual Policy Simulations (Excluding Kei-cars)

	Δe: Fuel consumption (liter/100km)	Δa: Weight (kg)	ΔL Welfare loss ($/car)	ΔL relative to efficient	S.D. of MC ($/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 4, 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.4: Counterfactual Policy Simulations (With Firm Dummy Variables Interacted with Cost Variables)

	Δe: Fuel consumption (liter/100km)	Δa: Weight (kg)	ΔL Welfare loss ($/car)	ΔL relative to efficient	S.D. of MC ($/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 4, 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.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).

Table A.6: 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 3 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.

References

Anderson, Michael and Maximilian Auffhammer. 2014. "Pounds that Kill: The External Costs of Vehicle Weight." *Review of Economic Studies* 82 (2):535–571.

Becker, Randy and Vernon Henderson. 2000. "Effects of Air Quality Regulations on Polluting Industries." *Journal of Political Economy* 108 (2):379–421.

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

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

Gruenspecht, Howard K. 1982. "Differentiated Regulation: The Case of Auto Emissions Standards." *American Economic Review* 72 (2):328–331.

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

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.

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.

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

Slemrod, Joel. 2010. "Buenas Notches: Lines and Notches in Tax System Design." Manuscript: University of Michigan.

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.

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.