Do Regulatory Loopholes Distort Technical Change? Evidence from New Vehicle Launches under the Japanese Fuel Economy Regulation

Yoshifumi Konishi†
Department of Economics
Keio University

and

Shunsuke Managi
Urban Institute and Department of Civil Engineering
Kyushu University

Forthcoming at *Journal of Environmental Economics and Management*

†Corresponding author. Mailing address: 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan. E-mail: ykonishi@econ.keio.ac.jp. Phone: +81-3-5427-1725.

Acknowledgement: Our special thanks go to Kong Joo Shin who provided valuable assistance in preparing the data for this manuscript. We also thank two anonymous reviewers, Hendrik Wolff (the co-editor), Werner Antweiler, Brian Copeland, Corbett Grainger, Sumeet Gulati, Tomohiko Inui, Daiji Kawaguchi, Carol McAusland, Hikaru Ogawa, Hiroshi Ohashi, and other seminar participants at the University of British Columbia, the University of Tokyo, the National Institute of Environmental Studies (Japan), the National Graduate Institute for Policy Studies, and the Research Institute of Economy, Trade, and Industry (Japan) for their helpful comments. The study was in part supported by the financial support from Japan Society for the Promotion of Science (JSPS), Grant-in-aid for Scientific Research (Grant number: 18K01562). This is a substantially improved version of the July 2019 draft.

Do Regulatory Loopholes Distort Technical Change?
Evidence from New Vehicle Launches under the Japanese Fuel Economy Regulation

Abstract: Environmental regulation often creates regulatory loopholes that are not ideal in first-best settings. Such loopholes affect the marginal costs of alternative compliance strategies, leading to distortion in firm's compliance choice. We quantify the unintended effect of such loopholes on technical change in the Japanese automobile industry, using variant-level data on new vehicle launches. We employ a triple difference strategy, exploiting the two-fold treatment-control structures within each product segment, due to regulation-induced variations in the Japanese fuel economy standards over time. Our results indicate that regulation-induced differences in technical trade-offs have induced a distortion not only in product attributes but also in technical progress in fuel economy technology.

JEL Codes: D22, K32, L62, Q48, Q55
Key Words: Automobile, triple difference, energy policy, fuel economy regulation, Ratchet effect, regulatory loopholes, technical change, technology policy

1. Introduction

Since Porter (1991), economists have long been interested in empirically examining the effect of environmental regulation on technical change [e.g., Newell, Jaffe, and Stavins (1999); Popp (2002); and Calel and Dechezleprêtr (2016)]. The literature to date, however, has primarily focused on the *direct* (or *intended*) effect of environmental regulation: i.e., the effect of a regulation-induced increase in the implicit price of pollution on technical change in sectors that use pollution as a factor of production [Copeland and Taylor (1994)]. Real-world environmental regulation, however, often entails design features that offer 'loopholes' that may not be necessarily ideal in first-best settings [Anderson and Sallee (2011), Sallee and Slemrod (2012), and Ito and Sallee (2018)]. Such design features may alter the marginal costs of available compliance strategies, thereby creating misguided incentives for firm's technology choice, the effect of which can potentially persist over time via the market size effect of technical progress [Acemoglu (2002), Acemoglu *et al.* (2012), and Aghion *et al.* (2016)].

Earlier studies find clear and convincing evidence that firms indeed exploit regulatory loopholes in a variety of contexts — a flexible-fuel exemption under the U.S. CAFE regulation (Anderson and Sallee, 2011), a notched schedule of the U.S. Gas Guzzler tax (Sallee and Slemrod, 2012), and weight-basing under the Japanese fuel-economy regulation (Ito and Sallee, 2018). These studies, however, primarily focus on firm's 'second-stage' product choice conditional on their 'first-stage' technology choice, leaving out potential distortion in technical change (Sallee and Slemrod, 2012). Consequently, these studies tend to imply that regulatory loopholes can be welfare-enhancing, offering regulated firms low-cost compliance strategies, given the second-best regulatory setups.[1] Such conclusion may change, however, if the distortion in the second stage choice also leads to the distortion in the first stage choice.

We investigate this question empirically in the context of automobile fuel-economy regulation in Japan. Doing so requires a model of firm behavior that accounts for both its choice over product attributes and that over technology investments. To that end, we build upon Knittel's (2011) insight that automakers face technical trade-offs between fuel economy and other vehicle attributes and that these trade-offs change persistently over time. This insight is incorporated into a simple model of firms, which allows us to (i) distinguish the distortion on the first-stage choice on technical capital versus that on the second-stage choice on product attributes, (ii) clarify how the two types of distortion arise through a simple economic mechanism, (iii) explain how the two types of distortion can be empirically identified, and

[1] Regulatory loopholes are, of course, not optimal in the first-best setting where efficient regulation can be costlessly implemented.

(iv) establish a clear, testable economic prediction: A fuel-economy regulation distorts technical change when it creates *any* trade-offs between the targeted and secondary attributes. The economic intuition is simple — attribute-based standards influence the relative marginal costs of two compliance strategies, compliance by attribute-shifting or compliance by technology adoption.

Our empirical strategy relies on the unique quasi-experimental setup created due to the Japanese weight-based fuel economy regulation. Under the regulation, the fuel economy standards are a step function (or notched function) of vehicle weight: i.e., vehicles are classified into discrete weight segments with varying levels of fuel economy standards. Importantly, when revising the standards in 2007, the regulatory authority chose narrower weight segments, effectively creating two or more weight bins within each old weight segment. Consequently, we have substantial variations, in terms of stringency and width, across weight bins over time. We translate these variations into two measures of regulatory assignment: the 'stringency' of fuel-economy standards, measured in relative terms to the old standards, and the 'slope' of fuel-economy regulation, measured as a decrease in the fuel economy standard per unit of weight increase. The latter is indeed a convenient measure of the attributes trade-offs induced by the fuel economy regulation. In principle then, we should be able to test our hypothesis by comparing the outcomes of car models assigned to different regulatory slopes. The key here is how to control for confounds that may be correlated with the regulatory assignments.

To do so, we combine a set of control strategies with a triple difference (DDD) research design. First, as shown in Knittel (2011), a vehicle's fuel economy is a function of other product attributes such as horsepower, torque, and most importantly, vehicle weight. Hence, by including these attributes as direct controls, we can identify the changes in firm's technical trade-offs, isolating the effects of up-weighting or manipulation of other product attributes (we provide a more in-depth discussion on this point in **Section 3**). Second, we employ a triple difference estimator, exploiting the three-fold control structures: (a) cross-sectional with models assigned to low regulatory slopes as a control group, (b) temporal with the pre-2007 period as a control period, and (c) within-group cross-sectional with models assigned to non-stringent standards as an additional control. This last control is particularly important if firms engage in the Ratchet-type behavior (we discuss this in more depth in **Section 5**).

We strengthen our triple-difference estimator by a few more strategies to control for time-varying confounds. First, we construct treatment-control pairs within each old weight segment, and use maker- and segment-fixed effects interacted with time dummies. This allows us to control for confounds that arise from (segment-level) consumer demand, firm-level heterogeneity in technical progress, and tax/subsidy incentives offered during the post-2007 period. Second, we use two alternative measures, one exploiting bin-level variations and another exploiting model-level variations, for both the 'slope' and the 'stringency' of the regulation. This way, we are able to attribute the difference in outcomes solely to the difference in attribute trade-offs created due to the fuel-economy regulation. We implement this DDD strategy using vehicle characteristics data for all domestic passenger vehicles introduced between 2004 and 2012, excluding electric, diesel, and hybrid cars as well as those launched in the interim regulatory period 2007-2009.

We have two important empirical findings. First, we find strong evidence that regulatory loopholes in this context had a sizable, statistically significant distortionary effect on technical change. Our DDD estimate indicates that a one-unit increase in the steepness of the regulation causes a 13-19 percentage point (ppt) reduction in the rate of fuel-economy improvement. We emphasize that we obtain this qualitatively large impact, isolating the effects of up-weighting as well as other time-varying confounds that likely correlate with regulatory stringency. The economic significance of this impact can be cast in light of the work by Knittel (2011). Using variant-level data from the U.S. automobile industry, Knittel estimates that U.S. passenger cars could have improved fuel economy by roughly 60% over the 25-year period between 1980 and 2006 if their curb weights (and other attributes) had stayed at the 1980 level. Employing a similar exercise, the Japanese passenger cars had the same rate of technical progress just over the 8-year period between 2004 and 2012. The 19-ppt reduction represents roughly 1/3 of this technical change. Our finding also substantiates the economic significance of earlier empirical findings (e.g., Anderson and Sallee, 2011; Sallee and Slemrod, 2012; and Ito and Sallee, 2018) as it provides evidence that the distortion in the second-stage attribute choice can lead to the distortion in the first-stage technology choice.

Second, we find that the notched schedule of the regulation does offer regulatory loopholes, but firm's incentives to exploit the loopholes vary substantially due to the differences in attribute trade-offs that arise from these notches. In other words, not every notch is equally important. We confirm this by examining the effect of regulatory slope on vehicle weights in a manner analogous to the main regression. Our results indicate that holding the stringency of the regulation, firms increase the weights of their vehicle models more when they are faced with steeper regulatory slopes than faced with flatter slopes. The magnitude of the impact is also qualitatively large — a one-unit increase in the steepness of the regulation causes a 12 ppt increase in vehicle weight. This finding also confirms the economic mechanism underlying our first finding.

Our study is related to several vibrant areas of research: (i) regulatory loopholes in environmental regulation (e.g., Anderson and Sallee, 2011; Sallee and Slemrod, 2012; and Ito and Sallee, 2018), (ii) the optimal design of regulation on transport-related emissions under second-best settings [see a review by Anderson *et al.* (2011) or Knittel (2012)], (iii) quantitative evaluation of Corporate Average Fuel Economy (CAFE) regulation (e.g., Austin and Dinan, 2005; Jacobsen, 2013; Goldberg, 1998); and (iv) the effect of environmental regulation on innovation and technical change [see a comprehensive review by Jaffe *et al.* (2002)]. Our findings have important implications for these strands of literature: the welfare cost of regulatory loopholes can be potentially larger than in the previous studies; subtle design features may matter for the efficiency properties as well as the economic evaluation of CAFE regulation under second-best settings since they affect the implicit cost of regulation; and regulatory loopholes can affect the rate of technical change in ways that environmental regulation is not originally intended. Recent advances in the empirical industrial organization literature indeed substantiate the importance of endogenous product/technology choice (e.g., Seim, 2006; Hitsch, 2006, Fan, 2013, Crawford *et al.*, 2015, Wollmann, 2018), both qualitatively and quantitatively, for policy and welfare evaluation. Our results suggest that incorporating this aspect in the empirical study of the fuel-economy regulation can be an important direction for future research.

The paper proceeds as follows. The next section describes the regulatory background. Section 3 sets up our empirical/theoretical framework, incorporating the concept of technology possibility frontiers into Ito-Sallee's analytical framework. Section 4 explains our data set. Section 5 discusses our identification and estimation strategy. The results are discussed in Section 6. The last section concludes, with a short discussion on the implications of our empirical findings for welfare and policy evaluation.

2. Regulatory Background

The Japanese fuel economy regulation is based on what is known as the *Top-runner* system. The system was first introduced under the 1999 Amendments to the Energy Conservation Act for all manufacturing products that consume energy in utilization. Under the *Top-runner* system, the government first classifies each vehicle to a unique product category according to its vehicle weight, and then chooses the highest observed fuel economy rating as the standard for that product category. This results in the fuel economy standards that are a step function of curb weights. The first weight-based fuel economy standards under this system were adopted in 2001 with a target year 2010. Since then, the standards were revised twice, in 2007 and 2013. Like the Corporate Average Fuel Economy (CAFE) standard in U.S., the Japanese fuel economy standards are enforced only at the firm level, based on the sales-weighted corporate average. **Figure 1** depicts the 2001 standards and the 2007 standards.[2] The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) adopted a new fuel economy rating method, known as *JC08 Mode*, for the new standards. The figure reports the new standards in the old measure (known as *10.15 Mode*). The method of conversion between the two measures is described in **Section 4** in more detail.

The new 2007 standards created an interesting regulatory setup, and thus, is a focus of our study. Under the new standards, the government chose a narrower weight segment to define each product category, resulting in 16 new weight segments in contrast to 9 under the old standards. As a result, each old weight segment was effectively divided into two or more bins, resulting effectively in 24 weight bins in total under the new standards. For some reason (not transparent in regulatory documents), the segment width differed substantially across weight segments. Furthermore, because the fuel economy performance of the top-runner relative to the peers in the same old weight segment differed substantially across different weight bins, the required fuel economy improvement relative to the old standard also differed substantially across these bins. Consequently, there are bins that are relatively 'steeper' than others relative to the old standards (the steepness or 'slope' is measured as a decrease in the fuel economy standard per unit of weight increase [a more detailed discussion on this point appears in **Section 5**]). We expect that this variability in slope and stringency levels across weight bins distorts economic incentives for firms' product offerings.

There are a few more regulatory backgrounds that become important in our empirical analysis. First, the Japanese fuel-economy regulation does not permit credit trading across firms. On one hand, this helps our identification since it eliminates the potential confounding effect from credit trading. On the other hand, it suggests the need to control for heterogeneous firm incentives because the marginal costs of compliance are unlikely to be equalized across firms. **Figure A1 in the online appendix** shows that at the beginning of the new standards, all domestic car makers were behind the required fuel economy standards, and hence, heterogeneity is probably more important at the car-model level. Second, the Japanese government introduced a series of tax/subsidy incentives since 2009. Interestingly, these incentives were tied to the 2001 standards, rather than the 2007 standards, until 2012 [for details, refer to Konishi and Meng (2017)]. Hence, firms faced the same tax incentives within each old weight segment until 2012. To isolate the confounding impact of these tax incentives, we make use of a treatment-control structure within each old segment, and also constrain our main empirical analysis up to year 2012 (see our identification and estimation strategy in **Section 5**).

[2] In this paper, we refer to the old standards as the "2001 standards" and the new standards as the "2007 standards" both for clarity and for economizing space, although they are often referred to as the 2010 and the 2015 standards, respectively, in the Japanese regulatory context.

Lastly, one more aspect of the Japanese regulation needs some discussion. Fines for non-compliance are only 1 million JPY ($\approx \$10,000$) per *firm*. Moreover, the Japanese standards are not enforced every year, and instead, firms are expected to meet the standards only by the (respective) target years. In contrast, under the U.S. CAFE, fines are $55 *per vehicle sold, for every mile-per-gallon shortfall*. The National Highway Traffic Safety Administration reports that the U.S. automobile industry has been paying roughly $20 million annually since 2010 (*AutomotiveNews*, July 16, 2016). Despite this weak incentive structure, however, Japanese firms take these standards very seriously, plausibly in fear of non-pecuniary sanctions such as damaging customer reputation and unfavorable treatment in public procurement. The firms met the 2001 standards in *every weight segment* by 2005 well ahead of its target year 2010 (and before the start of the tax/subsidy incentives in 2009). Hence, the new standards were adopted in July 2007. The firms again met the 2007 standards by 2012 before its target year 2015. Hence, the Japanese government again adopted the latest standards in March 2013 with a target year 2020. A recent scandal revealed that Mitsubishi Motors had been inflating fuel economy ratings for nearly 20 models over the last 10 years (*Japan Times*, Jun 17, 2016). Furthermore, Ito and Sallee (2018) show that firms do respond, very sharply indeed, to the weight cutoffs of the 2007 fuel economy standards (despite the fact that model-level tax/subsidy incentives are tied to the 2001 standards). These incidents seem to suggest that compliance with the standards (hence, non-pecuniary sanction for non-compliance) is indeed very costly for firms.

3. Empirical Framework

3.A. Theory of Attribute-based Regulation Revisited

Ito and Sallee (2018) present a theory of attribute-based regulation in an empirical context similar to ours. They call a technology regulation 'attribute-based' if it relies on a secondary attribute that is not the direct target of the regulation. Energy efficiency regulations around the world are often attribute-based. For example, fuel economy or carbon emissions standards are a function of vehicle footprint in the U.S. and of vehicle weight in Japan and the EU. Energy efficiency labels and standards for buildings, consumer electronics, and home appliances have similar features. Attribute-based regulations are often preferred over uniform regulations in the regulatory arena for efficiency as well as equity concerns. In this context, Ito and Sallee demonstrate that (1) in the presence of (efficient) credit trading, no attribute-basing (i.e., a flat standard) is optimal, but (2) some attribute-basing (i.e., a sloped standard) is optimal in its absence. Most importantly, their model clarifies that it is *not* optimal to perfectly equalize the marginal costs of compliance, highlighting the importance of striking a balance between marginal cost harmonization versus bias minimization in firm's attribute choice.

We extend their framework by introducing the notion of a technology possibility frontier (TPF) in the attribute space. TPF is defined as the set of product attributes that are technically feasible when technical inputs are used most efficiently given the technology capital. The concept is implicit in Ito and Sallee and other related studies, but in our view, has not been given proper attention. We believe that such TPFs do exist in the automobile industry on the basis of Knittel (2011), who finds that technical trade-offs exist between fuel economy and other vehicle attributes for automobiles in the U.S. market and that the technical trade-offs change over time as firms' technologies improve over time. We see a similar, remarkable shift in the technical trade-offs in the Japanese automobile industry over the last 25 years (see **Figure 2**, which displays technical trade-offs between fuel economy and curb weight for Toyota's passenger vehicles offered between 1990 and 2015).[3]

This concept of TPF helps us *conceptually* distinguish the impact of the first-stage technology choice versus that of the second-stage choice on product attributes. This distinction is empirically quite important, as Sallee and Slemrod (2012) write, in examining the impact of the notched schedule of the U.S. Gas Guzzler Tax, that their estimate gets at the welfare effect of marginally adjusting fuel economy ratings conditional on the "choices regarding engine size, body style and vehicle features that cannot be changed quickly and have large impacts on fuel economy" (p. 991). Furthermore, the concept also helps us *empirically* isolate the effect of the second-stage bias (which we observe directly) from that of the first-stage bias (which we do not observe). In this sense, the TPF concept allows us to naturally extend the work of Ito and Sallee (2018) to a study of distortion in technical change.

Below, we present a simple model of firm's choice over technology and product attributes under attribute-based regulation that explicitly incorporates the concept of TPF. The model not only helps us clarify what we do in this paper, but also provides a general framework for empirical analysis, with which to contrast and evaluate ours with other related studies [e.g., Ito and Sallee (2018), Jacobsen (2013), Klier and Linn (2015), Reynaert (2015), Whitefoot *et al.* (2017)]. In particular, the model is intended to demonstrate three essential points in a simple and unified framework: (i) a regulation-induced distortion in firm's choice over

[3] Note that we are *not* claiming here that the solid lines in **Figure 2** are the TPFs. We are simply saying that this regularity suggests that firms must be facing some TPFs, which we *do not* directly observe. In 1990, the (unweighted) average fuel economy of all Japanese passenger cars was roughly 13.1 km/L. In 2015, that number increased by more than 70% to 22.3 km/L. This improvement in fuel economy did not come from downsizing vehicle weight. Indeed, the average curb weight increased by roughly 10% from 1,169 kg in 1990 to 1,293 kg in 2015.

product attributes can also lead to a distortion in the technology investment, (ii) the distortionary incentives depend on the 'slope' of the regulatory constraint (to be clarified below), and (iii) differences in the regulatory slopes across product segments can induce differences in the levels of technology investment even if the regulatory constraint induces a single implicit price of regulation (which would be true, for example, if the compliance is based on a firm level average such as in the case of CAFE regulation).

Consider a single firm producing a product. Let this firm be the 'representative firm', which also acts as an agent that solves for the equilibrium of a competitive market that maximizes the economic surplus. The firm offers this product in a unique product segment (which is exogenously fixed). The economic logic presented here is essentially intact under an alternative model with firms producing more than one product as long as we make analogous regularity conditions. The product is described by two-dimensional product attributes (f, w). Let f represent a targeted attribute and w (a composite of) non-targeted attributes. For ease of interpretation, we call f 'fuel economy' and w 'vehicle weight', but with an understanding that changes in other attributes (such as driving performance, size, torque) are also implicit in the latter variable. The firm faces a two-period decision: Choose the next-period product attributes (f_1, w_1) given the current-period product attributes (f_0, w_0). All of the economic rents that result from the current-period choice are treated as 'sunk' at the time of choosing next-period attributes.

Given this setup, the representative firm's decision making proceeds in three stages.[4] In the first stage, the firm chooses the level of investment in technical capital $s \geq 0$, which shifts up the technology possibility frontier defined as:

$$f = T(w, s).$$

In the second stage (at the beginning of the next period), it chooses a profile of product attributes (f, w), fully anticipating the consumer demand. Then in the third stage, it sets the price at the marginal cost of production, and the quantity supplied is pinned down by the market equilibrium, which maximizes the economic surplus given (f, w). We simplify our analysis by denoting this third-stage economic surplus by $U(f, w)$. Note that the marginal cost of producing the product with attributes (f, w) is already part of U.

With no regulation, the firm chooses (the next-period) product attributes (f, w) in the

[4] Alternatively, we may formulate the firm's decision as one in which all choices (s, f, and w) are made simultaneously, and the results would be essentially identical. We clarify this point in the **online appendix**.

second stage, so as to maximize:

$$\max_{f,w,s\geq 0} \; U(f,w) - C(s), \qquad \text{subject to} \quad f \leq T(w,s), \tag{1}$$

where C is the fixed cost of investment, which is sunk at the time of choosing product attributes. The current-period technology capital is normalized to zero, so the level of investment is conveniently identified with the next-period technology capital s.

A few clarifying comments are in order.[5] First, the existing literature has identified a number of compliance strategies firms can take in response to a fuel-economy regulation: (a) sales-mixing (i.e., adjusting prices to shift sales to vehicles that meet the standards), (b) attribute-shifting (i.e., adjusting vehicle attributes such as performance, vehicle footprint, weight), (c) technology adoption (i.e., adoption of technologies that improve on fuel economy), and (d) trading of compliance credits across segments. Our model explicitly accounts for (b), (c), and (d), but not (a). Reynaert (2015), in contrast, explicitly considers (a), (c), and (d), but not (b). As discussed in Reynaert, however, (a) and (b) work very similarly in terms of their welfare consequences. Second, it is widely known in the literature [see Jaffe *et al.* (2002)] that technical change has three distinct phases: invention, innovation, and adoption/diffusion. In our empirical analysis, we have no means to distinguish these three types of technical change. Hence, investment in s in our model may constitute *any* type of actions for acquiring the technological capital. For example, the firm may already know fuel-saving technologies that are widely available in the market, and may simply deploy some of them into a vehicle instead of developing its own. Third, the model incorporates the impact of technology adoption on *both* the marginal cost *and* the fixed cost of production. To acquire technologies (either by invention or by adoption), the firm pays some fixed investment cost. Given the level of technology s, then the firm may exploit it fully or compromise on other attributes, moving along the TPF. This choice of attributes (f, w) affects the marginal cost of production, which in turn affects $U(f, w)$ in conjunction with the third-stage choice on product price. This formulation is consistent with other related studies [e.g., Ito and Sallee (2018), Reynaert (2015), and Whitefoot *et al.* (2017)].

Given this economic environment, the regulator imposes an attribute-based regulation R, which mandates $f \geq R(w)$ in stage 'zero' before the firm engages in this three-stage decision. The regulation is enforced at the firm level, which induces an implicit price of regulation μ. The firm faces the implicit price μ even when its fuel economy exceeds the standard. This is true because it can sell the credits to other firms or to other segments of its own, which is treated as an 'outside option' to close the model. Hence, with regulation, we write the firm's

[5] We thank a referee for helping us clarify the following points.

optimization as:

$$\max_{f,w,s\geq 0} \quad U(f,w) - C(s) - \mu(R(w) - f), \qquad \text{subject to} \quad f \leq T(w,s). \tag{2}$$

As another benchmark, we also consider a social planner's problem. Let σ be the marginal external benefit of f, which we assume is constant and is the inverse of the marginal external damage of emissions. In a perfectly competitive environment, the price equals the marginal cost of production, and the market equilibrium pins down the quantity supplied, maximizing the economic surplus given (f, w). Hence, the social valuation U of (f, w) coincides with the representative firm's valuation, and hence, the firm's optimal choice of attributes, price, and quantity coincides with the planner's in the absence of externality. Given this setup, the planner's optimization program is written as:

$$\max_{f,w,s\geq 0} \quad U(f,w) - C(s) + \sigma f, \qquad \text{subject to} \quad f \leq T(w,s). \tag{3}$$

Our interest lies in how the choice of (f, w, s) differs across regimes (1), (2), and (3).

To focus on the essentials, we make further simplifying assumptions on C, U, T and R. (A1) The TPF function T is linear, strictly increasing in s, and strictly decreasing in w, with $T_{ws} \equiv \partial^2 T/\partial w \partial s = 0$, in the neighborhood of (f_0, w_0). That is, technical upgrade can only shift up the linear TPF schedule in the relevant decision space. (A2) Welfare (or economic surplus) function U is increasing in f and w and twice-differentiable, and the iso-surplus curves are strictly convex in (f, w). (A3) The cost of technical upgrade C is increasing in s at an increasing rate. (A4) R is linear and $R_w \equiv dR/dw \leq 0$.

The linearity of T (and R) is not as restrictive as it may appear. As shown below [and in Knittel (2011)], linear regression is surprisingly well fit to observed attributes in logged values. Moreover, automakers can generally choose new car models around the neighborhoods of their pre-existing models given their platform designs, and are likely to face approximately linear technical trade-offs in the neighborhoods. Assumptions (A2) and (A3) are standard regularity conditions, which ensure the optimization program is well-behaved.

Under these conditions, we have the following proposition, the proof of which is available in the appendix:

Proposition: Under the competitive product and credit market,

(i) the optimal policy sets $R_w = 0$: i.e., no attribute-basing (Ito and Sallee, 2018);

(ii) given the technical capital s, the firm's choice of secondary attributes under attribute-based regulation $w^R(s,\alpha)$ increases with α, where α is defined as:

$$\alpha \equiv T_w - R_w,$$

and hence, we have $w^S(s) < w^N(s) < w^R(s,\alpha)$ if $\alpha > 0$ (where the superscript S stands for social optimum, N for no regulation, and R for regulation);

(iii) the attribute-based regulation (i.e., $R_w \neq 0$) always distorts the level of technical upgrade, and if $U_{fw} < 0$, the level of technical upgrade decreases with α.

Part (i) of the proposition replicates Proposition 1 of Ito and Sallee (2018). Our economic environment is virtually identical to theirs, except that we explicitly account for the effect of technical upgrade on the TPF. Hence, this result confirms the generality of their result. Part (ii) of the proposition demonstrates how the distortion in the second-stage attribute choice is related to the 'slope' of regulation. Part (iii) establishes that this distortion in the attribute space can also lead to the distortion in the first-stage choice on technology.

A few remarks are in order. First, in a sense, part (iii) of the proposition re-establishes what is already known in the literature: Attribute-shifting and technology adoption are substitutes, the firm's equilibrium choice equates the marginal costs of alternative compliance strategies, and regulatory loopholes can influence these marginal costs in an important way (Anderson and Sallee, 2011). What is new here, however, is that the slope of regulation acts as a loophole, and affects the relative marginal costs of these compliance strategies, in a way that differs from the social valuation. In the **online appendix,** we establish this connection between our results and the principle of marginal cost equalization across compliance strategies. Second, we refer to this regulation-induced bias in the level of technical capital as 'distortion' in technical change. It is important to note, however, that the firm's choice of s^R under the sloped regulation may still coincide with the social optimum s^S if the regulation induces a shadow price that exceeds the social marginal benefit σ. Hence, the word 'distortion' in this paper does not necessarily mean 'distortion relative to the optimum'. Lastly, the condition $U_{fw} < 0$ may arise quite naturally in our empirical context either because the marginal consumer valuation of fuel economy decreases with vehicle performance and size (e.g., Berry *et al.*, 1995) or because the marginal cost of improving fuel economy increases with these attributes.[6]

[6] Our proposition establishes $U_{fw} < 0$ as one sufficient condition for the lower level of technical investment. However, the bias in the second-stage choice on attributes can lead to slower technical progress more generally because the suboptimal expansion path for $(f(s), w(s))$ under the attribute-based regulation can decrease marginal returns to investment for other reasons.

The essence of our results is best explained graphically. Imagine a product segment, say 'light sedan', which implicitly defines a (small) two-dimensional segment on the attributes space (f, w). On this segment, the current product offering, denoted O in **Figure 3**, is on the solid linear line, which represents the current-period TPF. The dashed curves represent firm's iso-surplus curves (IC) in the next period. Without the regulation, the firm chooses a new product attribute (f, w), labeled A, at the tangency between the iso-surplus curve and the next period's TPF. Hence, the optimal bundle is uniquely pinned down given the level of technical upgrade s. The firm then chooses the level of technical upgrade s such that the marginal cost of doing so equals the marginal increase in profits.

Let us first consider the impact of regulation on product attributes. Consider first the optimal (i.e., no attribute-basing) regulation. The regulation induces a shadow price on fuel economy, which equals μ. Hence, it achieves the optimum as long as $\mu = \sigma$. With no attribute-basing, and holding the technology level constant, the firm seeks to minimize the cost of compliance by moving *up* on the TPF (i.e., to the left). Hence, the private optimum occurs at the tangency, labeled B, between the TPF and the *flatter* iso-surplus curve (i.e., the dashed green line). This curve indeed coincides with the social iso-surplus curve (SIC) on (f, w) if $\mu = \sigma$.

What happens under attribute-basing? To see, let the solid red line, denoted R, represent the attribute-based regulation. For ease of exposition, we draw the case where R is steeper than T and cuts through T from the above. The first thing to note is that this significantly affects the firm's marginal incentive. Given the technology level, moving *up* along the TPF to the left of A increases the implicit tax payment while moving *down* to the right of A decreases it. Hence, it undermines the incentive to use down-sizing of vehicles as a means to improve fuel economy. Note that making the slope of regulation the same as the TPF does not completely eliminate this incentive — it is still 'sloped' relative to the optimal (no attribute-basing) regulation, and hence, it undermines the incentive. When the slope of regulation is sufficiently steep, the firm's private optimum occurs at the tangency, such as C, between the TPF and the *steeper* iso-surplus curve.

Let us examine the impact of attribute-basing on technical upgrade. Recall, first, that the dashed green curve represents the social iso-surplus curve on (f, w). Let's draw another indifference curve that goes through C. The social planner would not pick such a bundle because it attains a lower social welfare than bundle A at the same cost of technical upgrade. This lowers the firm's incentive to invest in s. Because the firm equates the marginal cost of technical upgrade with the lower marginal benefit, an optimum bundle must lie on a lower TPF, like D. Indeed, the same reasoning explains why the investment is lower under no regulation than under regulation.

The above exposition also helps us clarify a few other points. First, the figure helps us distinguish the two types of distortion: the one on technology choice (from B to B') and the other on product attribute holding the technology level constant (from B' to D). Second, it also demonstrates that the incentive to manipulate on the secondary attribute (i.e., up-weighting) is directly tied to the incentive to invest in technical capital on the targeted attribute (i.e., fuel economy technology). Hence, the distortion in the second-stage choice can lead to the distortion in the first-stage choice. Third, this distortionary incentive depends on the slope of regulation relative to that of TPF. Lastly, the figure also points to a challenge when we attempt to empirically distinguish the two types of distortion. As is clear from the figure, neither the *direction* nor the *magnitude* of change in vehicle weight tells us the direction or the magnitude of bias in technical change. Compare D versus E, for instance. E represents a larger increase in weight relative to O than D does. Yet, D lies a lower TPF than E, implying a larger technical bias. The figure, however, also hints us a way to overcome this challenge. By definition, 'technical change' is directly associated with an expansion of TPF. Assuming the rationality of firm's decision, product bundles must lie on the TPF. Then, if we find two vehicle offerings identical in all attributes (e.g., such as brand, horsepower, weight), yet differ in fuel economy ratings (in a statistical sense), we can infer this difference as the difference in technology level. Indeed, earlier studies [e.g., Knittel (2011), Newell *et al.* (1999); Popp (2002)] rely on the same underlying principle for identification of technical progress.

3.B. Empirical Model

Our empirical approach is primarily data-driven. Japan's fuel economy standards are set at the model-variant level whereas their enforcement is based on sales-weighted averages at the manufacturer level. Since the regulation is enforced at the manufacturer level, we would ideally model manufacturers' strategic incentives to offer different variants of different car models in different years explicitly, fully endogenizing both pricing and product choice (e.g., Seim, 2006; Hitsch, 2006, Fan, 2013, Crawford *et al.*, 2015, Wollmann, 2018). However, such structural modeling of endogenous product choice requires demand-side information that is far more detailed than we have at hand. Since we are interested in the effect of the fuel economy standards that are imposed at a variant level, we need demand-side information that can vary at a variant level. With more than 1,000 variants offered each year, we lack enough sources of variation to separately identify the influences of variant-level demand factors from those of the regulation in the structural framework.

We thus take a simpler approach, and focus on the reduced-form estimate of the impact of the standards on firms' technology possibility frontiers, exploiting policy-induced variations across weight segments over time in a triple difference (DDD) research design. To that end, we follow Knittel (2011) (in spirit) and define TPF as follows. Fuel economy f of vehicle variant i is a function of a vector of observable product attributes $\mathbf{x}$ (incl. vehicle weight), and a variable s that expresses the level of technology capital. We also posit that the technology-augmenting component is multiplicatively separable:

$$f_i = F(\mathbf{x}_i, s_i) = \phi(s_i)G(\mathbf{x}_i).$$

In other words, we define 'technical change' in fuel economy as the change in the level curve connecting the set of product attributes ('inputs') that would produce the same fuel economy ('output'). This assumption is arguably restrictive, but a similar assumption has been used widely in other related studies. For example, both micro- and macro-level studies on total factor productivity often rely on Cobb-Douglas or CES production functions. Empirical regularities found in Knittel (2011) and our data also support the validity of this assumption. Given a combustion engine type (i.e., diesel, electric, fuel, and hybrid), the technical attributes trade-offs seem rather stable over time — the curves that represent the technical trade-offs show persistent patterns over time, with only changes in the level of the curves over time.

Given this assumption, the empirical object of interest is given, in logged form, by:

$$\ln f_i = \ln \phi(s_i) + \ln G(\mathbf{x}_i) + \epsilon_i. \tag{4}$$

Our identification and estimation strategy exploits several advantages stemming from this specification (see **Section 5**). First, the first term, which captures the level of technical capital, is additively separable from the second term. Hence, the distortionary impact of regulation on technical capital is also separable, falling only on the first term, after controlling for the second term. Second, it also allows us to isolate the distortionary impact of regulation on technical change from those that directly arise from distortion on secondary attributes (such as weight and horsepower). As evident from Ito and Sallee (2018), the Japanese fuel economy regulation induced firms to up-weight their vehicles. This up-weighting directly decreases their fuel economy ratings even in the absence of technical change. However, eq. (4) tells us that we can purge out this effect by directly controlling for the second term — the effect of up-weighting on fuel economy must come through the second term if it does not influence technical change.

This empirical model relies on a few more assumptions for identification. First, we assume that technology capital exists at the car-model level. Firms often delegate development of a car model to a specific group of engineers in the form of a division or a team, and the group of engineers apply and accumulate knowledge and technology in designing the car. Hence, firm's technology frontier can vary at the model level, at least in the short run. Note that our notion of technology capital is broadly defined here (and consistent with the common usage in the economics literature). As discussed in Knittel (2011), a number of fuel-saving technologies are widely available in the market. Such technologies include combustion improvements, low resistance tires, reduced drivetrain friction, transmission improvements, turbocharger/supercharger, and variable valve timing. Firms need not to necessarily 'develop' these, and instead, may simply 'deploy' them in order to improve fuel economy. Doing so on a specific vehicle model is, by definition, an addition to the model-specific technology capital — even the same know-how and technology must be tailored and adjusted to a specific model. We emphasize here this assumption neither imply nor require that technology capital does not exist at the firm level or segment level. All we require is the existence of some technology capital at the model level.[7] Given this nature of technical progress, we posit that firms choose the level of model-specific technology capital in response to model-level regulatory assignment. For example, if a firm sees that many variants of a car model fall in a very tight fuel economy standard, then it makes variant-level choices for that car model in the subsequent period. Such a firm may decide to eliminate all grades for the car model entirely, change the combinations of fuel-saving technologies, re-design the platform, or offer a completely new model under a different name.

Second, the framework presented in **Subsection 3.A.** tells us that attribute-basing of any degree leads to distortion in product choice and technical change. In other words, the distortion occurs *both* when the regulatory slope is higher *and* lower than the TPF slope. This poses a challenge in identifying the regulatory impact because we do not observe the TPF in the absence of regulation, and hence, we cannot directly compare the regulatory slope with the TPF slope. In the case of automobiles, however, it is known to be extremely costly for the firm to decrease weight given the vehicle's platform design. Hence, in the present context, we assume that the distortionary incentives are unidirectional. Hence, the hypothesis to be tested in our empirical context is, *The weight-based fuel economy regulation distorts technical change if the regulatory slope is higher than the slope of the (average) firm's TPF.* This unidirectional nature helps us use policy-induced variations for identifying the distortionary impact of the regulation.

[7] Nowadays, it is very common for automakers to share technologies and platform designs across different models. Hence, technology capital does exist at a higher level than the model level. However, there is still likely to be a difference between the level of technology capital at the model level versus that at the firm or the shared-model level. That difference is all that is required for our empirical strategy.

Under these assumptions, we should be able to identify the impact of the fuel economy regulation, in principle, by comparing differences in $\ln \phi(s_{jmt})$ in eq. (4) across different vehicle models assigned to different weight segments. The challenge, of course, is how to control for other confounds that might have affected the TPFs. We discuss our identification and estimation strategy in more detail in **Section 5**.

4. Data

Our data come from Carsensor.com, one of the largest online car retailers in Japan. The compiled data set contains variant-level information on observable attributes of virtually all vehicles sold since 1991: e.g., model year/month, curb weight, displacement level, fuel economy rating, horsepower, list price, size, torque, transmission and other available options. Importantly, because we have information on grade year/month at the variant level, we can identify the year in which each vehicle variant was first offered to the market. Our main analysis covers a subsample vehicles launched during the 2004-2012 excluding observations in 2007-2009 because the new standards are implemented in July 2007 and we anticipate that it takes at least a few years before the regulation influences firm's technical capital. Hence, we use 2004-2006 as the pre-treatment control period and 2010-2012 as the treatment period. More detailed justifications for this choice follow below.[8] In our placebo analysis, we also use observations from 2001 to 2003 and from 2013 to 2015.

We drop diesel, electric, and hybrid vehicles as well as commercial vehicles since they are not subject to the same fuel economy regulation as outlined in **Section 2.**[9] We also drop observations on imported brands because foreign manufacturers can always choose to sell a subset of their models to Japan, and thus, their TPFs are unlikely to fully respond to the incentives created through the Japanese regulation. We also exclude vehicles produced by Mitsubishi Motors because it might severely contaminate our results if included, since the recent scandal revealed that their reported fuel economy ratings during our study period do not follow the same regulatory guidelines as others.

A complication arises in compiling fuel economy data. The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) changed the method to measure fuel economy as an

[8] The statistical significance and direction of the regulatory impact are largely intact, though the magnitude of the impact does change, if we also include observations from 2008 and 2009.

[9] There is a separate weight-based fuel efficiency regulation on diesel cars. The sales of diesel cars accounts for a tiny portion of the overall sales in the Japanese market. Hence, to avoid unduly complications, we drop diesel cars from our analysis. Hybrid vehicles are subject to the same regulation, but their fuel economy ratings are well above the fuel economy standards, and therefore, weight category assignment should not influence their technical progress.

effort to align reported fuel economy with actual on-road fuel economy. As a result, all new vehicles offered after October 2010 must report fuel economy in a new measure, known as *JC08 Mode*, while all vehicles offered before October 2010 report in an old measure, known as *10.15 Mode*. These two measures are not directly comparable. Fortunately, however, the MLIT also mandated that all old vehicles must also record fuel economy in *JC08 Mode* if they are still sold in the market. Hence, the Japanese manufactures tend to report fuel economy in both measures in our study period. We fit a regression of 10.15-mode fuel economy on JC08-mode fuel economy on these observations, and then use the predicted fuel economy in case of vehicles missing fuel economy data in *10.15 Mode*.[10] From here on, all fuel economy data are reported in the *10.15 mode*.

We clarify important differences between our data and the data used in Ito and Sallee (2018). Ito-Sallee data come from the list of new cars published each year by the MLIT. The MLIT list reports data at the car configuration (or '*Katashiki*') level, which is coarser than the grade level reported in the Carsensor catalog. There are two aspects of the MLIT data that make it unsuitable for our analysis. First, the MLIT list contains all cars sold as 'new cars' as of the end of each fiscal year. As a result, some cars are reported in multiple years in the MLIT data. For example, Toyota Vitz 2010-model, which was sold as a new car between December 2010 and April 2012, are reported twice in fiscal years 2010 and 2011. Our data do not suffer from this double counting because we have information on model years and we count each observation only once for the year it was first launched. Second, the MLIT list reports only the *range* of vehicle weights for about 3/4 of the reported car models while reporting the *unique* fuel-economy rating for each model. This range can be as large as 200 kg, averaging at around 35 kg. Hence, the MLIT data would imply each car model has a flat TPF in the short run. In contrast, every model variant in our catalog data reports a unique value of curb weight (and of fuel economy). The **online appendix** provides a more detailed discussion on these points.

5. Identification and Estimation

5.A. Identification Strategy: An Overview

To motivate our empirical strategy, we re-write the equation (4) into the following parametric form.

$$\ln f_{ijmst} = \mathbf{A}_{ijmst} + \mathbf{X}'_{ijt}\gamma + \epsilon_{ijmst}, \tag{5}$$

[10] The regression is surprisingly well fit with $R^2 \approx 0.99$.

where $\ln f_{ijmst}$ is a logged fuel economy of vehicle variant i of model j, in weight segment s, introduced by firm m in year t, $\mathbf{A}_{ijmst}$ represents the terms that capture the level of technical capital, and $\mathbf{X}_{ijt}$ is a vector of observable vehicle characteristics [incl. weight (w), horsepower (hp), size ($size$), torque (tq) (all in logged values), and transmission type]. The first term $\mathbf{A}$ is intended to capture $\ln \phi(s)$ whereas the second term $\mathbf{X}'\boldsymbol{\gamma}$ is meant to directly control $\ln G(\mathbf{x})$ in eq. (4).

The goal of our empirical study is to estimate the causal (distortionary) impact on $\mathbf{A}$ of the technical attribute trade-offs imposed by the fuel-economy regulation. To do so, we need to control for confounds that may be correlated with regulatory assignment — i.e., confounds that remain after controlling for observable vehicle characteristics such as weight, horsepower, and manufacturer fixed effects. The first is the consumer demand. It is known that consumers who buy larger and heavier cars tend to care less about fuel economy ratings (Berry *et al.*, 1995; Konishi and Zhao, 2017). Car models under different weight categories may naturally have different rates of technical progress, irrespective of regulatory assignment. Second, as discussed in **Section 2**, the government offered tax/subsidy incentives according to fuel economy ratings since 2009. We thus need to tease out the effect of these tax incentives. The third, and probably the most important, confounder concerns the Ratchet effect.[11] The Ratchet effect refers to the phenomenon that the agent under-performs to avoid a demanding schedule in the future in a dynamic incentive scheme where the principal updates the scheme over time upon observing the agent's performance (e.g., Freixas *et al.*, 1985; Laffont and Tirole, 1988). If exists, the Ratchet effect would imply that the rate of technical change in a weight category in the future may correlate with the current fuel-economy standard for that category since under the *Top-runner system*, the regulator chooses the best observed fuel-economy rating as the standard. It is difficult to isolate the Ratchet effect from the effect of regulatory trade-offs because both arise from the same design features of the fuel-economy regulation.

We attempt to control all these confounds in the following ways. First, we create treatment-control pairs within each old weight segment, exploiting the fact that the 2007 standards created new and narrower weight categories.[12] In addition, we include weight-segment dummies, interacted with time-series dummies, in our regression analysis. By this, we are able to compare the outcomes of regulatory assignments for car models that faced

[11] We thank Hiroshi Ohashi for suggesting this point.

[12] The government offered eco-car subsidy and tax credits based on fuel economy improvements relative to the old 2001 standards, despite that the new 2007 standards were already in effect (see **Section 2**). As we shall discuss below, we construct our treatment-control pairs within each 2001 weight segment, constructing regulatory slopes based on the new 2007 standards. In the **online appendix**, we discuss how firms might have manipulated in reporting their car model weights to the government, and the reported weights clearly responded to the 2001 standards, not the 2007 standards, during the 2010-2012 period.

roughly the same demand shocks and the same tax incentives over time. Second, we also include maker fixed effects, again interacted with the time dummies, to control for firm-specific technical progress. Third, we exploit two types of variations in regulatory assignment, i.e., changes in the level of the fuel-economy standards and changes in the width of the weight segments. As discussed in **Section 2**, some weight segments under the 2007 standard are more stringent than others in terms of required improvements relative to the old standards. At the same time, some segments are narrower than others, resulting in variations in the width of weight segments. We transform these variations into two measures of regulatory assignment: (1) the 'stringency' of fuel-economy standards, measured in relative terms to the old standards, and (2) the 'slope' of fuel-economy standards, measured as a decrease in fuel-economy standards per unit of decrease in vehicle weight. For robustness, we construct two alternative variables for each of these measures. These measures, we hope, would get at two types of economic incentives separately, the Ratchet effect and the effect of attribute trade-offs (we call it the 'slope effect' henceforth). We discuss these measures in more depth in the next subsection. Lastly, we combine these with a triple difference (DDD) approach, exploiting the three-fold control structures as follows:

(a) Cross-sectional between-group variation in regulatory slope

(b) Temporal variation over years (with years 2004-2006 as control)

(c) Cross-sectional within-group variation in regulatory stringency

Here, we treat car models faced with the same (or similar) regulatory slope(s) as a group. By using temporal variation with years 2004-2006 as an additional control, we are able to control for any stationary differences across groups as well as time-varying factors that are common to the groups. However, this difference-in-differences (DD) structure is not sufficient to control for the Ratchet effect or other time-varying confounds that affect these groups differently over time. To take care of this concern, we use another within-group variation. As we shall discuss more below, the stringency of the standards most likely captures the Ratchet behavior, and hence, serves as an additional within-group control. That is, we compare the outcomes of vehicles assigned to different regulatory slopes, but with the same (or similar) regulatory stringency level(s). The resulting DDD estimate is consistent under a weaker identifying assumption: i.e., *unobservables that affect the rate of technical progress differently across car models assigned to different regulatory slopes do not systematically differ across car models assigned to different stringency levels.* Besides the weaker condition for identification, this DDD structure comes with an additional benefit.

That is, any pairwise DD estimate, in addition to the DDD estimate, is also consistent if any pair of treatment/control groups satisfies the standard common-trend assumption. For example, if the contemporaneous shocks that affected the high-slope and low-cost groups have the same trend over time, then the DD estimate on a subsample consisting only of the same stringency level is also consistent.

Nesting all these control strategies in eq. (5), we arrive at the following equation for estimation.

$$\begin{aligned} \ln f_{ijmst} \;=\;& \alpha + \beta_1 R_t + \beta_2 T_j + \beta_3 H_j \\ & ... + \beta_4 (R_t \times T_j) + \beta_5 (R_t \times H_j) + \beta_6 (T_j \times H_j) \\ & ... + \beta_7 (R_t \times T_j \times H_j) + \mathbf{X}'_{ijt} \boldsymbol{\gamma} + \boldsymbol{\eta}_{mst} + \epsilon_{ijmst}, \end{aligned} \tag{6}$$

where R_t indexes a regulatory period and equals 1 during the post-2007 period, T_j and H_j are our key treatment variables ('slope' and 'stringency' of the 2007 standards, respectively), $\mathbf{X}_{ijt}$ is a vector of observable attributes defined in eq. (5), and $\boldsymbol{\eta}_{mst}$ denotes maker- and segment-fixed effects and their interactions with R_t. Note that $\mathbf{A}_{ijmst}$ in eq. (4) is replaced by the sum of two terms in eq. (5). The first is the triple-difference terms [in a manner analogous to Gruber (1994)], which is intended to identify the effect of regulation-induced distortion on technical change. The second is the controls for the influence of unobservables (i.e., $\boldsymbol{\eta}_{mst}$) on technical change. We estimate (6) using the variant-level catalog data, with alternative measures of both T_j and H_j to be discussed in the next subsection. The OLS estimate of β_7 identifies the causal impact of the regulatory slope under a much weaker assumption than the common trend assumption. One (potential) disadvantage of our empirical approach is that our specification assumes the regulatory assignment can affect only the *level* of the technical frontier, not the *slope*, over time. As Knittel (2011) points out, the estimates of technical progress (and, hence, the DDD estimate) may be biased downward if the technical trade-offs between fuel economy and other attributes are not as large in later years.[13] Of course, one could always allow slope coefficients to vary, say, by interacting them with our treatment

[13] There is an alternative strategy. That is, to use propensity score matching to control for the effects of these observable covariates. A disadvantage of the PSM estimator is that it requires a stronger identifying assumption than the DDD regression. That is, the DDD regression only requires that conditional on a set of covariates, differences in trend for unobservables between the treatment and the control groups stay the same between the high-cost and the low-cost segments while the PSM requires that the unobservables have zero means conditional on the set of covariates (i.e., conditional independence assumption). Because the PSM does not control for differences in unobservable time trend between the treatment and the control groups, the PSM estimates may be biased upward if the control groups exhibit a larger change in unobservable factors. Because the PSM is likely biased upward and the DDD is likely biased downward, the true impact of the regulation is likely to fall somewhere inbetween. Our earlier attempt to employ PSM estimator confirms this prediction.

variables. However, our empirical strategies primarily exploit *within-segment* variations, and we doubt that we have large enough within variations to credibly identify the impact on the slope parameter(s).

4.B. Slope vs. Ratchet

To illustrate our empirical strategy, let us take an old weight segment 1,010-1,265 kg as an example for our exposition. In **Figure 4**, both old and new fuel economy standards are drawn (red and blue lines, respectively). The intersections of the two standards create three weight bins on this old weight segment. The first point to note is that under the *Top-runner* system, the government essentially chooses the highest fuel economy rating that was achieved for each weight bin as the standard for that bin. This means that the standards approximately trace out the technology frontier of 'the most fuel efficient' vehicles that were available as of 2007. Consider a line connecting the two endpoints A and B of this segment. For the moment, let this line represent the technical frontier of a 'typical' or average firm. Next, consider another line connecting the two endpoints A and C of the lightest weight bin on this segment. For the moment, let us call the slope of this line the 'regulatory slope' of this weight bin. Then this regulatory slope is clearly steeper than the technical frontier. Then by virtue of our discussion in **Section 3**, we should expect the firm to increase curb weight and to improve less in fuel economy for vehicles that lie in this weight bin. More generally, some bins have steeper slopes than others on a given weight segment, as shown in **Figure 4**. Consequently, we should expect vehicles assigned to the high-slope bins to lie on a lower technical frontier than those assigned to the low-slope bins in the future model changes. As discussed in **Subsection 3.A.**, the fact that the standards are enforced on sales-weighted averages simply accelerates this incentive to increase curb weight for vehicles in the high-slope bins because it is easier for the firms to meet the overall standards if the firms have more car variants in low-slope weight bins. Note that our argument does not quite depend on the assumption that the line connecting points A and B of the old segment represents the technical frontier of a 'typical' firm. What matters for our empirical analysis is that vehicle models that lie on the same old segment are likely to face, on average, roughly the same market demand, the same regulatory incentives other than the slope, and the same technology frontier prior to the 2007 standards.

The question then is, what would be the most appropriate measure of the regulatory trade-offs? We consider two alternatives. The first measure directly applies the above logic, and calculates the slope of each bin b as the slope of a line connecting the two endpoints of the weight bin:

$$T_b = \left| \frac{h_{b+1}^{\text{new}} - h_b^{\text{new}}}{w_{b+1}^{\text{new}} - w_b^{\text{new}}} \right|, \tag{7}$$

where w_b and h_b are, respectively, the weight cutoff and the fuel economy standard for bth weight bin under the new 2007 standards. For the high-slope bin in **Figure 4**, this measure is simply the slope of the line connecting A and C.[14] The advantage of this measure is that it uses only the variations in regulatory assignment, and hence, it is unlikely to be correlated with other confounders at the firm or the model level, especially after controlling for the Ratchet effect (which we discuss below).

The disadvantage, however, is that it fails to account for firm-level or model-level heterogeneity. Even within a weight bin, different car models have different fuel-economy ratings and vehicle weights at the onset of the new standards, and these differences in initial positioning are likely to present different regulatory trade-offs. For example, a car model located at position O would be able to lower the standard by Δh by increasing its weight by Δw. This reduction Δh represents a large gain relative to its required fuel-economy improvement. In contrast, another car model located at position O' would be able to attain the same benefit, but by increasing its weight more by $\Delta w'$. Our second measure, therefore, accounts for this heterogeneity arising from initial positioning. That is, we calculate the regulatory slope for each car model j as a reduction in the fuel-economy standard, expressed as a percentage of the required fuel-economy improvement for that model, per unit of weight increase required for that model:

$$T_j^* = \left| \frac{(h_{b+1}^{\text{new}} - h_b^{\text{new}})/(h_b^{\text{new}} - f_j)}{w_{b+1}^{\text{new}} - w_j} \right|. \tag{8}$$

Albeit its merit, the potential disadvantage of this measure is that it may be correlated with unobservables that affect the rate of technical progress, even after controlling for the Ratchet effect, because it explicitly uses initial fuel-economy information in its calculation.

We now turn to a more intricate confounder, the Ratchet effect. By construction, the regulatory slope depends on the stringency of the fuel-economy standards, and we have two reasons to believe that it can capture the Ratchet effect, rather than or in addition to, the slope effect. First, firms cannot observe competitors' technical progress prior to their product launches, and hence, they can only base their Ratchet strategy on the regulatory standards. Second, their future Ratchet behavior depends on their past Ratchet behavior. Because the fuel-economy standards are the outcome of the past Ratchet behavior, they can be directly associated with the future Ratchet behavior. This logic suggests that we should

[14] In the analysis below, we use raw values of fuel economy and weight to calculate T's and H's. Alternatively, we could use logged values. We present the results of our main regressions in the **online appendix**, and confirm that the results are qualitatively intact.

be able to control for the Ratchet effect by controlling for the stringency of the fuel-economy standards.[15] The key identifying assumption here is that firms can influence the level of fuel-economy standards by manipulating the rate of technical change, but cannot affect how the weight category is chosen, so the width (and the resulting slope) of each weight segment is an exogenous shock to the firms.

The remaining question is, what would be the appropriate measure of regulatory stringency? Like the slope effect, we consider two alternatives. The first measure simply computes the difference between the old standard and the new standard for each weight bin b:

$$H_b = h_b^{\text{new}} - h_b^{\text{old}}, \tag{9}$$

whereas the second measure computes the difference between the new standard and the pre-policy fuel-economy rating for each car model j:

$$H_j^* = h_b^{\text{new}} - f_j. \tag{10}$$

In other words, the first measure simply evaluates the absolute stringency of the new standard for each weight bin while the second measure evaluates the relative stringency for each car model. The pros and cons of these stringency measures are analogous to those of the two slope measures. Because the *Top-runner* system chooses the highest observed fuel-economy rating as the standard for that segment, the first measure is likely to be directly related only to the top performer's rate of technical progress. In contrast, the second measure is related to the own rate of progress (relative to the top performer). The latter closely captures the Ratchet-type incentives for each car model, but is more likely to be endogenous than the former.

We clarify two practical issues. First, our regulatory variables T_j and H_j vary at the car-model level, not the variant level. We can only do this because one variant introduced in a year cannot be credibly identified with another introduced in a different year.[16] Therefore, we trace out model histories, so that all vehicle variants introduced during the post-2007 period can be associated, via model identifiers, with those introduced during the pre-2007 period. For models that continue to exist, this is easy because they can be easily matched

[15] In the competitive environment like ours, firms may either ratchet up or down because a top-performer's behavior can not only affect its rivals' costs but also its own. Hence, we control the Rachet behavior indirectly by the stringency of the standards.

[16] It is highly questionable to identify vehicle variants according to their attribute data, at least in our context. For example, suppose we observe two variants of Honda Civic, one introduced in 2004 and another in 2012, that have the same displacement, horsepower, etc. Suppose further that Honda Civic went through a significant platform change between the two years — many models would indeed go through such model change during such a long period. In that case, it seems natural to treat these variants as different variants.

by model identifier. For discontinued models, we search through publicly available articles and company reports to see if there is any successor model for each retired model. What complicates the issue is that not all variants of a model necessarily fall in a single weight bin because there are many variants of each vehicle model. To address it, we calculate the unweighted mean of vehicle weights of all variants for each vehicle model during the pre-2007 period, and then classify the vehicle model according to that mean.[17] Second, as noted above, we would like to ensure a treatment-control pair in each of old weight segments. For the second measures, this can be easily achieved since there are several models, with sufficient variations, in every segment. The problem is with the first measures, which vary only by weight bin. To ensure treatment-control pairs in all segments, we classify weight bins into high- versus low-slope bins according to whether their slopes are steeper than the slope of the joint segment connecting all bins within each old segment as illustrated in **Figure 4**. By this, we are assuming that vehicles within each old weight segment faced roughly the same technical frontier and that a new segment steeper than this average slope provides more incentives to manipulate on vehicle weight. Similarly, we also classify weight bins into quartiles of stringency levels, with 1 denoting bins that fall in the lowest 25th percentile and 4 that fall in the highest 25th percentiles.

Table 1 clarifies these points, highlighting the main sources of variation we exploit in our analysis. Each row represents a weight bin, which we define as the intersection of the old and the new weight segments. The solid lines represent weight segments under the 2001 standards, and the dashed lines represent those of the 2007 standards. For each of these weight bins, we report old and new fuel-economy standards, regulatory slope and stringency in two alternative definitions, the number of vehicle variants, and the mean and standard deviation of fuel-economy ratings during the pre- and the post-2007 periods. The first measure of slope T [eq. (7)] calculates the slope using the 'height' and 'width' of weight bins, and therefore, has a unique value for each weight bin (recall **Figure 4**). The next column reports 1 if this slope is higher than the overall slope of the weight segment joining all weight bins that belong to the weight segment. The second measure T^* [eq. (8)] instead computes the slopes for all car models that belong to each bin, and therefore, we

[17] Assignment based only on a single year, say, 2006 or 2007, is problematic in our setup because each vehicle observation is recorded with the year in which that vehicle was first offered. Because Japanese car models typically run on a 3-4 year cycle, including all the three-year observations likely cover all variants of models that are still produced as of 2006.

Figure A4 in the **online appendix** reports the summary of model histories and box diagrams describing the distribution of variant-level curb weights for car models assigned to the high-slope weight bins. Of the 30 models, 11 models did not introduce any new variants between 2010 and 2012, and thus, are classified as 'discontinued'. Of these 11 models, only 2 models had clear successor models. Others either had no clear successor model or were merged to another existing model.

report the mean and standard deviation in each row. Similarly, we report two measures of regulatory stringency H [eq. (9)] and H^* [eq. (10)] in an analogous manner. There is no high-slope weight bin that falls in either the 1st quartile or the 3rd quartile of stringency levels. Hence, for cleaner results, we drop vehicle models that fall in the 1st and the 3rd stringency quartiles. This also eliminates bins that are too narrow to have any product offerings (i.e., rows 17 and 23). Once we remove these bins, we have substantial variations in both slope and stringency measures across weight bins. There is some indication that firms are avoiding new offerings in the high-cost/high-slope weight bins. This is really an analogue of the 'bunching' effect Ito and Sallee (2018) point out. However, the tendency is not necessarily clear — there are high-cost/high-slope weight bins that received roughly the same number of new offerings between the pre-2007 and the post-2007 periods. This occurs presumably because firms may strategically offer models in the stringent weight segments as a way to avoid tough competition in less stringent segments.

5.C. Descriptive Evidence

Before moving to our main analysis, we take a glance at graphical evidence. We first make use of weight-bin-level variations in regulatory assignment. **Panel (a) of Figure 5** displays an unconditional scatter plot of logged fuel-economy ratings against logged vehicle weights for vehicle grades introduced before the 2007 standards. The figure excludes imported cars, commercial vans and trucks, diesel, electric, and hybrid cars as well as vehicles that fall in the 1st and 3rd quartiles of stringency levels during the pre-2007 period. Variants of vehicle models assigned to the high-slope weight bins are marked with circle; those assigned to the low-slope bins are marked with $\times$ [Here, the definition of high- versus low-slope follows **column (5) in Table 1**]. The figure indicates no sign of a significant difference in the technical trade-offs between fuel economy and weight prior to the new standards.

Panel (b) of Figure 5 repeats the same for those introduced between 2010 and 2012 under the new standards. In this figure, variants of the successor models of those assigned to the high-slope bins are also marked with circle. We now see some difference in technical trade-offs between the two groups. However, the effect is somewhat ambiguous in this figure because firms also upgrade other vehicle attributes. For better visibility, we condition out the influence of vehicle attributes other than those that directly relate to vehicle weight. **Figures (c)-(d)** essentially are the same as those in **Figures (a)-(b)**, except that the former plot the residuals from a regression of logged fuel economy on key vehicle attributes (in logged values) after removing the linear projection from terms involving horsepower, torque, transmission, and brand dummies. We now see the technical trade-offs of those assigned to the high-slope bins lie *far below* those assigned to the low-slope bins after the new standards despite the fact that the former lie *slightly above* the latter before the new standards. This offers support for both our economic mechanism and our empirical strategy.

Next, we turn to the model-level variations in regulatory assignment. We first compute the (unweighted) means of fuel-economy ratings (over vehicle variants) for each model, for each of the pre- and the post-2007 periods. We then calculate the changes in these means between the two periods. **Figure 6** then plots these changes against the regulatory slopes that account for initial positioning at the model level [per equation (8)] for different levels of regulatory stringency. The figure indeed shows the patterns consistent with our economic predictions. Many of the models improved (average) fuel-economy ratings after the new standards. And these improvements are indeed greater for those faced with more demanding fuel-economy targets (relative to their initial positions). Yet, the improvements seem to decline, and turn even negative in some cases, with the increase in regulatory slope.

To offer support for the common-trend assumptions, we plot **(a)** the means of fuel economy ratings by year and by treatment (i.e., high-slope vs. low-slope groups) and **(b)** the differences in the mean fuel economy ratings between the high-stringency and the low-stringency groups by year by treatment in **Figure 7**. **Figure 7-(a)** demonstrates that both groups showed a steady increase in average fuel economy, yet the low-cost group increased fuel economy more sharply after 2009. The figure does seem to refute the concern that those assigned to the high-slope segments tend to be those that attained high rates of technical progress prior to the assignment. However, the temporal patterns between the two groups before 2007 do not appear quite identical, suggesting there might be other confounders that affect the two groups differently over time. In contrast, **Figure 7-(b)** demonstrates that the differences in average fuel economy between the high- and the low-stringency groups have roughly identical temporal patterns between the high-slope and the low-slope groups. This boosts our confidence in our DDD estimates. The figures also point to another complication we might take into account. They show that changes in responses to the regulatory assignment are more discernible after 2009, rather than immediately after the regulatory change in 2007. This may be attributed to the fact that it takes generally a few years for firms to introduce new vehicle variants to fully respond to the regulatory change or that firms' incentives to respond to the new standards became stronger after the old 2001 standards expired in 2009.

6. Results

6.A. DDD Regression

Table 2 reports the results of four regression models for each of the two alternative measures of regulatory assignment. **Panel A** displays the results using the bin-level regulatory variations in T and H. The first model (in columns 1 and 2) in this panel estimates DD regressions on the pooled sample, with high-slope bins against low-slope bins as the primary treatment. The estimates from these regressions would be biased downward if vehicles assigned to the low-stringency weight bins also respond to the high slopes, even if the common-trend assumption between the treated and the control groups is satisfied. The second and third models estimate the same regressions, but on subsamples consisting only of those of high-cost bins and of low-cost bins, respectively. The last model estimates full DDD regressions on the pooled sample. Each of these models is estimated with or without segment dummies interacted with time dummies. All specifications include weight (w), horsepower (hp), size ($size$), torque (tq) (all in logged values) and AT/CVT dummy as well as brand dummies interacted with time dummies. **Panel B** essentially repeats the same, but using the model-level regulatory variations T^* and H^*.

We first discuss **Panel A**. The DD estimates of the impact of the high slope on the pooled sample are negative but statistically insignificant. The magnitude of the estimates gets much larger and statistically significant when the same regressions are run on a subsample consisting only of high-stringency weight bins. In contrast, the DD estimates turn statistically insignificant on a subsample consisting only of those assigned to the low-cost bins. These results are consistent with our expectation, and are indeed suggestive of the success in our empirical strategies. Firms have a greater incentive to exploit regulatory loopholes when faced with more stringent standards. Per our theory presented in **Section 3**, this incentive result in a lower rate of progress in fuel-economy technology. However, the fact that the DD estimate on the low-stringency subsample is positive if time-varying segment controls are not included, but turns negative (and insignificant) once these controls are included implies that the rates of technical progress do vary across segments, irrespective of regulatory assignment, and are indeed higher for vehicles assigned to the low-stringency bins. This in turn suggests that vehicles assigned to the high-stringency bins might have been those with a lower rate of technical progress, and hence, if uncontrolled, this might confound the DD estimates since the regulatory slope correlates with the regulatory stringency. Hence, this gives support for our DDD strategy. The DDD estimate is indeed negative, statistically significant, and qualitatively very large: The estimate implies that the assignment to high-slope weight bins slows down fuel-economy improvements by roughly 13-19 ppt. Because we control for all relevant covariates, this also implies that the TPF for those assigned to high-slope weight bins would have lied strictly above the observed TPF if they had been assigned to low-slope weight bins instead. These results also explain why the observed TPFs seem flatter for those assigned to high-slope bins than those assigned to low-slope bins after than before the 2007 in **Figure 5-(b) or (d)**. As shown in **Table 1**, heavier weight bins tend to have less stringent standards (i.e., lower compliance costs). The assignment to high-slope bins in these heavier weight bins does not induce quantitatively large impacts on technical progress, whereas it has large negative impacts in lighter weight bins. Consequently, the observed TPF should look flatter.

Next, we turn to **Panel B**. The results here are quantitatively very similar to those in **Panel A**, but differ qualitatively on one important account. Recall that our regulatory variables T and H in this panel incorporate variations in vehicle models' initial positioning (prior to the 2007 standards) relative to the 2007 standards. Hence, there remains substantial variation in regulatory stringency H across vehicles within a subsample consisting only of either the high-stringency or the low-stringency bins. Therefore, if the Ratchet-type effect indeed exists and arises due to this model-level regulatory stringency, the DD regression estimates would be biased on all samples (i.e., in all of columns 1-6). On each subsample, the bin-level regulatory stringency is roughly controlled, and hence, much of the remaining variation arises from the variation in vehicle models' initial positioning. Since our slope measure is a direct function of this model-level regulatory stringency, vehicles faced with steeper (model-level) slopes may be simply those that had initially low fuel-economy ratings. The DD estimates on these subsamples, therefore, may be simply picking up the effect of having initially low fuel-economy ratings relative to the new standards. Though the direction of the bias is hard to predict a priori, our DD results seem to suggest a plausible direction. Our DD estimates are negative (and statistically significant) on the subsample consisting of the high-stringency bins whereas they are positive (and statistically significant) on the low-stringency bins. We may interpret this result as follows. An additional increase in regulatory stringency (at the model level) induces the Ratchet-type effect and slows down the rate of technical progress only when the regulatory stringency is already very high. The Ratchet-type incentive disappears, however, when the stringency level is low. Therefore, an additional increase in stringency simply leads to a higher rate of technical progress.[18] Our DDD strategy helps us control for this effect, giving us an unbiased estimate of the slope effect. The DDD estimate on the pooled sample is indeed negative and statistically significant. The magnitude is also large — a one-unit increase in the regulatory slope slows down fuel-economy improvements by roughly 17-28 ppt.

[18] In theory, firms face a uniform shadow price of fuel economy regulation across weight bins irrespective of their stringency levels as long as the regulation is enforced at the firm level. However, firms may have other non-pecuniary incentives to comply with the standard at the model level. For example, it is common for firms to display a commercial label on a car model, indicating its fuel economy performance relative to both old and new standards. Such model-level incentives are expected to affect the Ratchet-type behavior.

6.B. Economic Mechanism

Our results so far confirm a statistically and qualitatively large impact of regulatory assignment to high-slope weight bins. A question remains as to exactly what economic mechanism caused that effect. The economic mechanism outlined in **Section 3** is that weight bins that have steeper slopes relative to the pre-existing TPFs would induce firm to increase vehicle weights. If this is indeed the economic mechanism, we should also observe an increase in average curb weight for vehicle models assigned to the high-slope weight bins. Identifying this effect is, however, more intricate than identifying the effect on fuel economy for several reasons.

First, this logic suggests that vehicles in such weight bins should increase weights only up to the next weight cutoffs. This means that the anticipated weight increase should be bound, in principle, by bin size (measured as $|w_{b+1} - w_b|$ in kg). This is in contrast to fuel economy improvements, for which there is no apparent bound because how much to improve fuel economy given other product attributes (incl. weight) should only depend on the net marginal benefits of doing so. Hence, from the outset, the expected impact on curb weight may not be large enough compared to the variance of curb weight for each weight bin. This issue is further complicated by the fact that there is large variation in bin size. Larger (i.e., longer) weight bins may exhibit two counteractive effects. First, because firms have incentives to increase vehicle weight only to the next weight cutoffs, we might expect a larger weight increase in larger weight bins. However, larger weight bins also mean that it takes a more weight increase to cross the next weight cutoff. Given the design and size of a vehicle, it may be easy to increase weight by, say, 20 kg, but may be hard to increase weight by, say, 100 kg. A priori, there is no clear reason to expect which effect is stronger.

The reasoning suggests that for cleaner results, we might control for bin size. To do so, we first calculate bin sizes of all segments (excluding the lightest and the heaviest weight bins), and classify them into quartiles of bin sizes. By tabulating our main sample by these quartiles, we find that the 1st bin size quartile (i.e., the smallest bins) contains observations in all stringency $\times$ slope subsamples. Hence, we run DD and DDD regressions of logged curb weight on the same set of covariates as in **Table 2** (excluding logged weight, of course). **Table 3** reports the results of these regressions, for each of the two alternative measures of regulatory assignment as in **Table 2**.

In **Panel A**, the DD estimate is positive and statistically highly significant on the high-stringency weight bins, but is not significant on the low-stringency bins. These are consistent with our results on fuel-economy ratings. The DD estimate on the pooled sample averages out these two, and hence, is positive but statistically insignificant. On the other hand, the DDD estimate gets at the differences between the two, and hence, is positive and statistically highly significant. These results seem to confirm that high regulatory slopes do create incentives to increase curb weight, particularly in high-stringency bins. **Panel B** essentially confirms the same point. Recall that all DD estimates in **Panel B** are likely to give us biased estimates for essentially the same reason discussed in the previous subsection — the DD estimates may simply pick up the effect of having initially low fuel-economy ratings relative to the new standards rather than the effect of assignment to high slopes. Hence, we focus on the DDD estimate. There, we again see a large and statistically significant, positive effect of high-slope assignment on curb weight.

6.C. Placebo Checks

We run two placebo checks to verify our regression results. For ease of interpretation, these placebo experiments perturb only the bin-level regulatory assignments. Hence, our placebo results are directly comparable to those reported in **Panel A of Table 2**. First, we arbitrarily perturb our weight-bin assignment and see if our results continue to hold. Specifically, we shift weight cutoffs w_b in (7) by an arbitrary number k (in kg) and run the same regressions as before. The parameter estimates from this fictitious assignment should qualitatively differ from those of the factual weight-bin assignment. This placebo assignment, however, may not simply result in the disappearance of the statistical significance since we expect intricate influence of such placebo assignment in almost every weight bin. **Panel A of Table 4** below reports the DD and DDD estimates when $k = 25$ and all the same covariates as in **Table 2** (incl. brand and segment dummies interacted with time dummies) are used.[19] The DD estimates on the pooled sample and on the high-stringency subsample (columns 1 and 2) are statistically insignificant. Furthermore, the DD estimate on the low-stringency subsample turns negative and statistically highly significant. We take these as a support for our main results. The placebo experiments make the distortionary effect of the high-slope assignment go away on the samples where we expect it to be large, while making it stronger on the sample where we expect it to be small. Consequently, the DDD estimate is negative but statistically insignificant. This boosts credence in our main results.

Next, we perturb on temporal dimension, holding weight bin assignment. In **Panel B of Table 4**, we report the DD and DDD estimates, using 2003-2004 (instead of 2004-2006) as the

[19] Note that we cannot choose k to be too small or too large. Because we average weights over all variants of each model, virtually all models would be assigned to the same weight bins if we choose k to be too small. In the meantime, choosing too large a number is problematic because it would end by shifting virtually all models to the next weight bins. The average bin size is roughly 75 kg. Hence, we end up choosing a number between 20 and 30.

control period and 2005-2006 (instead of 2010-2012) as the fictitious treatment period (again, all the other covariates stay the same as in **Panel A of Table 2**). Because the new standards are adopted in July 2007, the estimates on this placebo treatment should not capture the effects of differential regulatory treatments due to the new standards. This placebo exercise should, instead, capture the pre-trends across different weight-bin assignments, and therefore, also serve as the check for the common-trend assumptions. The DD estimates on both the high-stringency and the low-stringency subsamples are positive and statistically insignificant at conventional levels. Accordingly, the DD estimate on the pooled sample is also positive and statistically insignificant. Furthermore, the DDD estimate on this fictitious treatment is negative and statistically insignificant. These results support our main results — we pass the common-trend checks (both with and without taking the third difference) and the fictitious treatment gives us the results that are qualitatively much different from the main results. These results also imply that those assigned to high-slope bins are associated, though not statistically significant, with slightly higher rates of technical progress during the pre-2007 period. This gives us another reason why we might prefer our triple-difference estimate over the DD estimate.

7. Concluding Remarks

Environmental regulation often creates regulatory loopholes that may not be ideal in first-best settings. We quantify the unintended effect of such loopholes on technical change in the context of Japanese fuel economy regulation, using variant-level data on new vehicle launches. We build upon the work of Knittel (2011), formalizing the notion of a technology possibility frontier in a simple, unified framework similar to Ito and Sallee (2018). The framework helps us conceptually distinguish the distortion on the first-stage choice on technical capital versus that on the second-stage choice on product attributes. It also helps us motivate how one may empirically distinguish the two types of distortion. Importantly, our model demonstrates that the 'slope' of attribute-based regulation works as a regulatory loophole and that such a loophole can lead to the distortion in technical capital, via a simple principle of marginal cost equalization, even in the presence of a uniform shadow price of regulation. In this sense, our framework re-casts the importance of Knittel's work in re-interpreting the work of Ito and Sallee (2018).

To test our economic prediction, we exploit quasi-experimental variations in the Japanese fuel economy regulation over time. Under the regulation, fuel economy standards are a step function of vehicle weight, and these standards change substantially over time, in terms of both stringency and width across weight bins. We exploit these policy-induced variations in a triple difference estimator. We then augment the estimator with a set of control strategies to tease out the effects of other time-varying confounds. Our results indicate that regulation-induced differences in technical trade-offs have indeed induced a distortion not only in product attributes but also in technical progress in fuel economy technology. In particular, our estimate suggests that assignment to high-slope weight bins slowed down the rate of fuel economy improvements (relative to low-slope weight bins) by roughly 13-19 ppt. We caution, however, that our estimates only get at the bias in technical progress that arises due to differences in regulatory slopes; hence, our results do *not* imply that the Japanese fuel economy regulation reduced the overall rate of technical progress, either relative to the social optimum or to the no-regulation counterfactual.

Our findings have two important implications for welfare and policy evaluation. First, the welfare cost of regulatory loopholes can be potentially larger than found in earlier studies. For example, Anderson and Sallee (2011) write, in their study on the flexible-fuel credits under the CAFE regulation, "the flexible-fuel loophole may actually increase welfare by allowing firms to relax an inefficient (fuel-economy standards) constraint (p. 106, parenthesis added). Such conclusion may change if it also slows down the technical progress in fuel economy technologies. Second, our results reinforce the importance of accounting for firm's technology choice in the optimal design of environmental regulation in second-best settings. Our results imply that, even when an efficient market for fuel-economy credits is in place, the attribute-based fuel economy standard can still bias firm's technology adoption because it can influence the marginal costs of alternative compliance strategies. Attribute-basing can naturally arise in other regulatory setups (e.g., carbon tax and feebates), and the bias in firm's technology adoption may also have a second-order impact on technology spillover and innovation. To what extent the competition in the market adds to the inefficiency loss from such a distortion is a priori uncertain, and therefore, can be an important agenda for future research.

References

[1] Acemoglu, D. (2002). Directed Technical Change. *Review of Economic Studies* 69(4): 781–809.

[2] Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The Environment and Directed Technical Change. *American Economic Review* 102 (1):131–166.

[3] Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. V. Reenen. (2016). Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry. *Journal of Political Economy* 124(1): 1-51.

[4] Anderson, S. T., and J. 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.

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

[6] Austin, D. and T. Dinan. (2005). Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes. *Journal of Environmental Economics and Management* 50: 562–582

[7] Berry, Steven, James Levinsohn, and Ariel Pakes. 1995 Automobile Prices in Market Equilibrium. *Econometrica* 63 (4): 841-890.

[8] Calel, Raphael and Antoine Dechezleprêtre. (2016). Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market. *The Review of Economics and Statistics* 98(1): 173–191.

[9] Copeland, Brian R. and M. Scott Taylor. (1994). North-South Trade and the Environment. *The Quarterly Journal of Economics* 109 (3): 755-787.

[10] Fan, Y. (2013). Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market. *American Economic Review* 103 (5): 1598-1628.

[11] Freixas, X., Guesnerie, R., & Tirole, J. (1985). Planning under Incomplete Information and the Ratchet Effect. *The Review of Economic Studies* 52 (2), 173-191.

[12] Goldberg, Pinelopi Koujianou. 1998. The Effects of the Corporate Average Fuel Efficiency Standards in the US. *Journal of Industrial Economics* 46 (1): 1-33.

[13] Goolsbee, A. and A. Petrin. (2004). The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV. 72 (2): 351-381.

[14] Gruber, J. (1994) The Incidence of Mandated Maternity Benefits. *American Economic Review* 84(3): 622-641.

[15] Hitsch, G. J. (2006). An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty. *Marketing Science* 25 (1): 25–50.

[16] Ito, K. and J. Sallee (2018). The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards. *The Review of Economics and Statistics* 100 (2): 319-336.

[17] Jacobsen, Mark R. (2013). Evaluating US Fuel Economy Standards in a Model with Producer and Household Heterogeneity. *American Economic Journal: Economic Policy* 5 (2): 148-87.

[18] Jaffe, A.B., Newell, R.G. and Stavins, R.N. (2002). Environmental Policy and Technological Change. *Environmental and Resource Economics* 22: 41-70.

[19] Klier, T. and J. Linn. (2012). New-vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard. *RAND Journal of Economics* 43 (1): 186-213.

[20] Knittel, C. R. (2012). Reducing Petroleum Consumption from Transportation. *Journal of Economic Perspectives* 26 (1): 93-118.

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

[22] Konishi, Yoshifumi and Meng Zhao. (2017). Can Green Car Taxes Restore Efficiency? Evidence from the Japanese New Car Market. *Journal of the Association of Environmental and Resource Economists* 4(1): 51-87.

[23] Laffont, Jean-Jacques and Tirole, Jean. (1988). The Dynamics of Incentive Contracts. *Econometrica* 56 (5): 1153-75.

[24] Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The Induced Innovation Hypothesis and Energy-Saving Technological Change. *The Quarterly Journal of Economics* 114(3): 941–975.

[25] Popp, D. (2002). Induced Innovation and Energy Prices. *The American Economic Review* 92 (1): 160–180.

[26] Porter, M. (1991). America's Green Strategy. *Scientific American* 264(4), 168.

[27] Reynaert, M. (2015). Abatement Strategies and the Cost of Environmental Regulation: Emission Standards on the European Car Market. *Working paper.*

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

[29] Seim, K. (2006). An Empirical Model of Firm Entry with Endogenous Product-type Choices. *RAND Journal of Economics* 37 (3): 619-640.

[30] Whitefoot, K., M. Fowlie, and S. Skerlos. (2017). Compliance by Design: Influence of Acceleration Trade-offs on CO2 Emissions and Costs of Fuel Economy and Greenhouse Gas Regulations. *Environmental Science and Technology* 51 (18): 10307-10315.

[31] Wollmann, T. (2018). Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicle. *American Economic Review* 108 (6): 1364-1406.

Figure 1. The Old and New Fuel Economy Standards

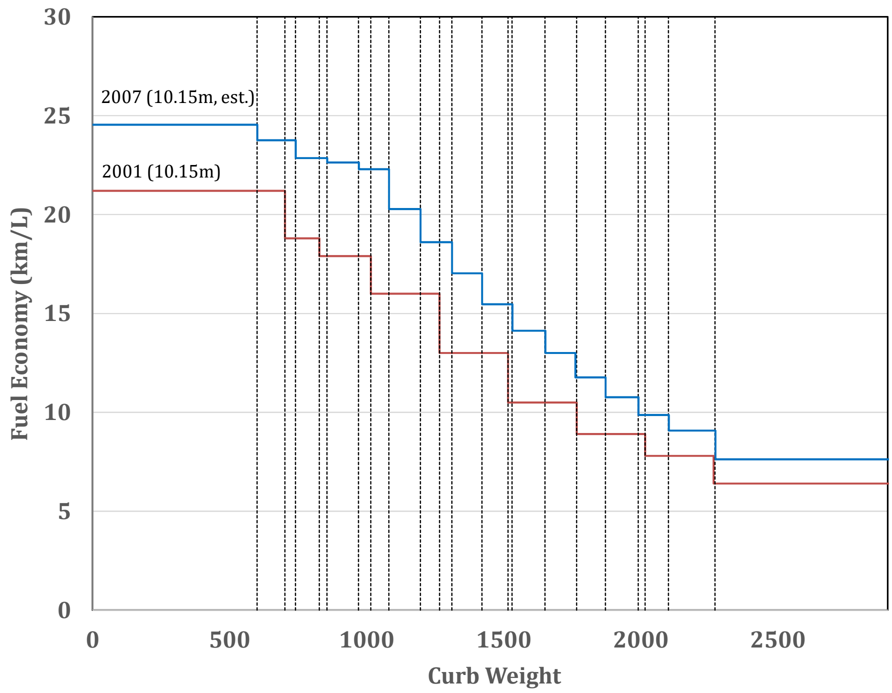
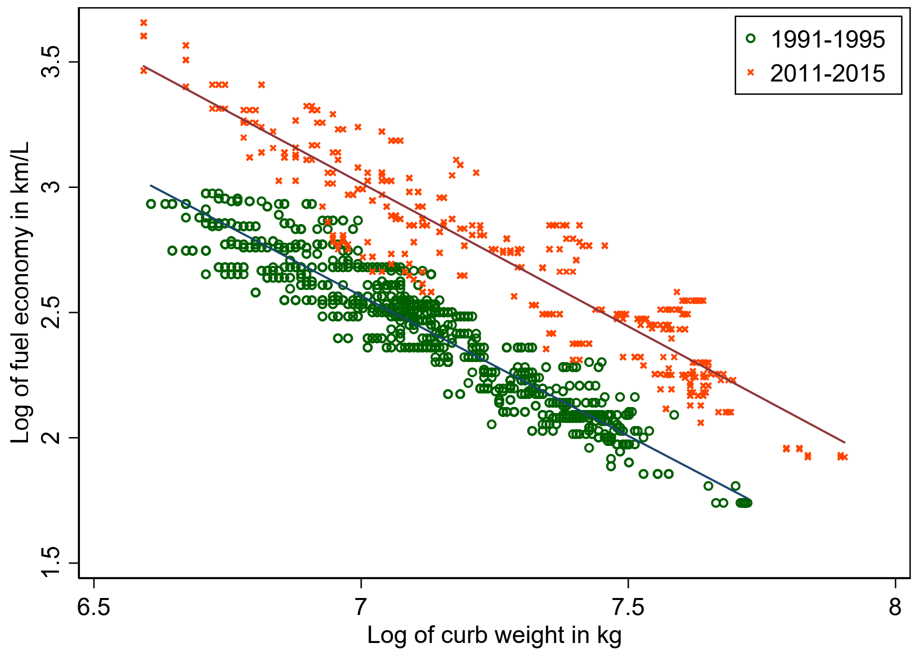


Figure 2. Changes in Technology Trade-offs for Toyota's Passenger Cars between 1991 and 2015



Note: The figure excludes commercial vans and trucks, imported brands, diesel, hybrid, and electric cars.

Figure 3. Impact of Attribute-based Regulation on TPF

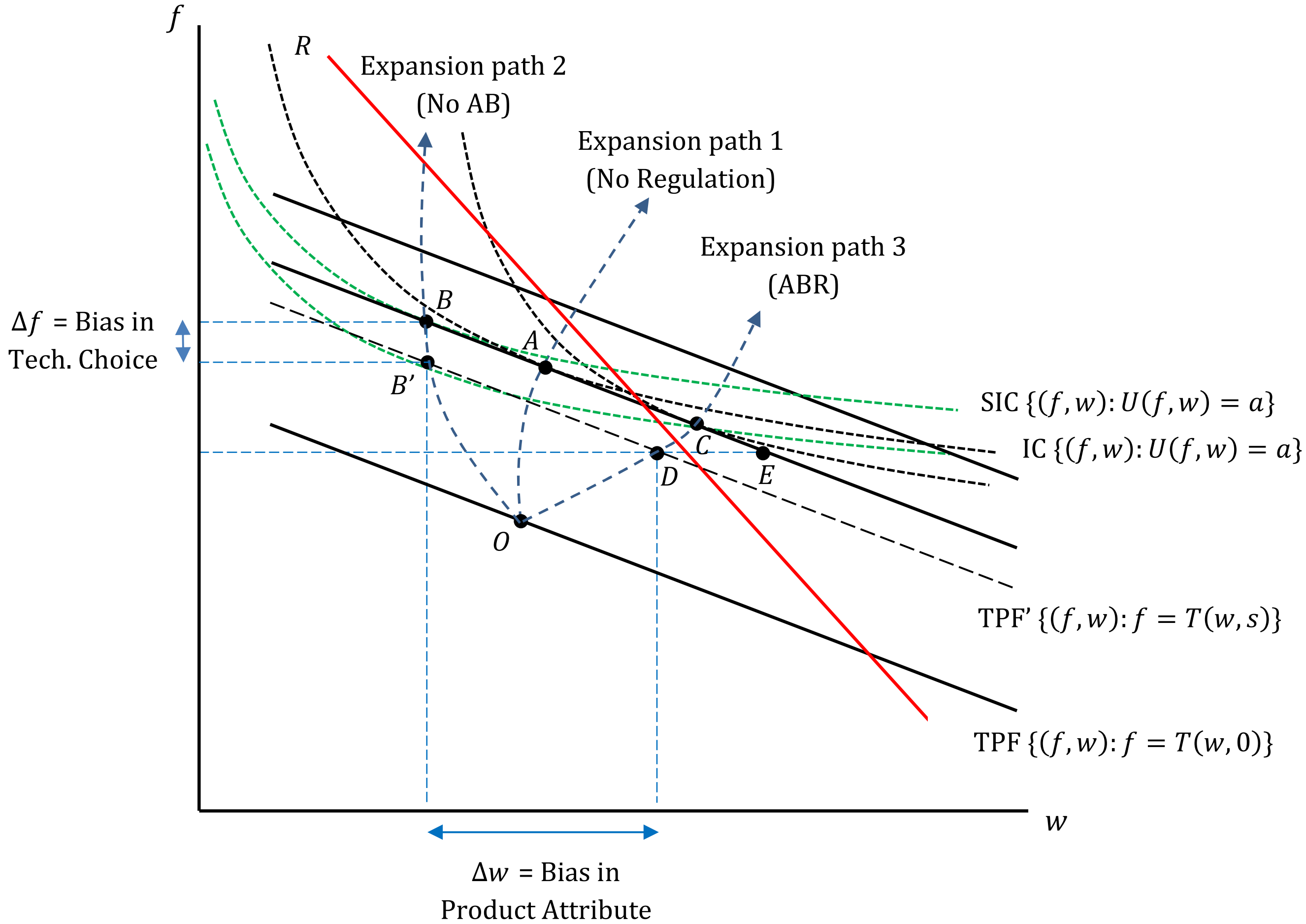


Figure 4. Variation in Regulatory Assignments: An Illustration

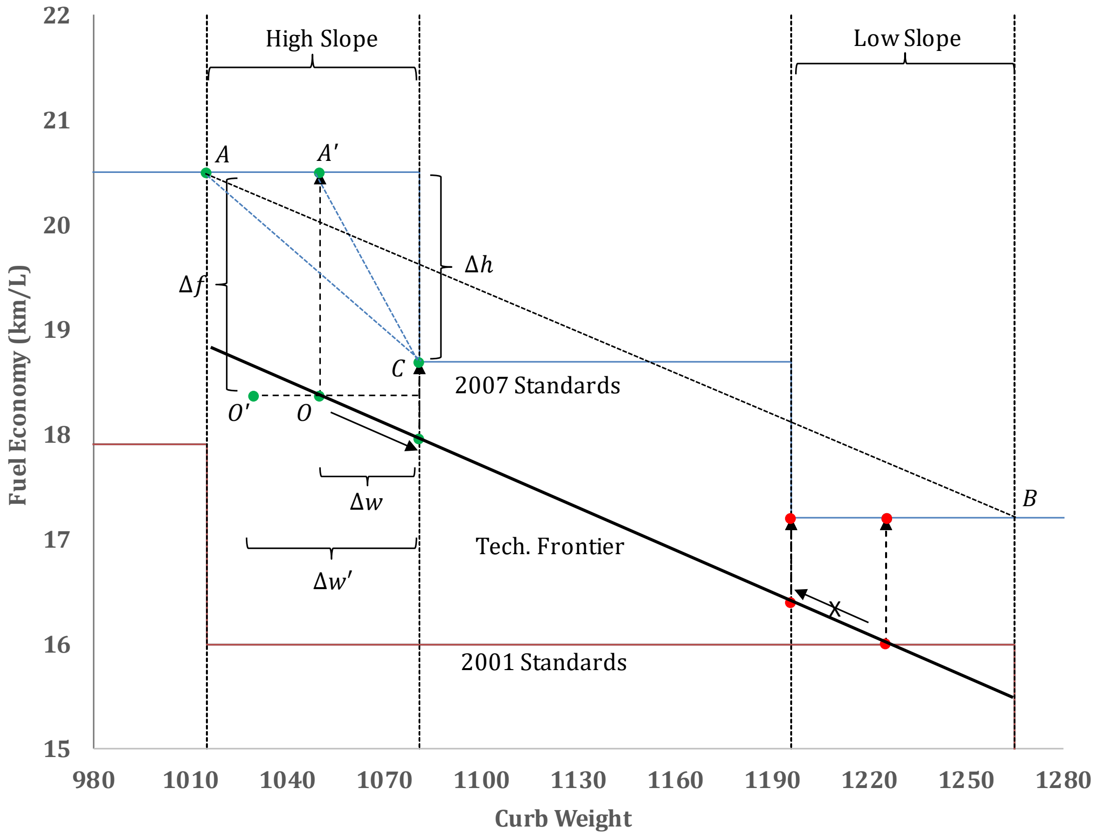
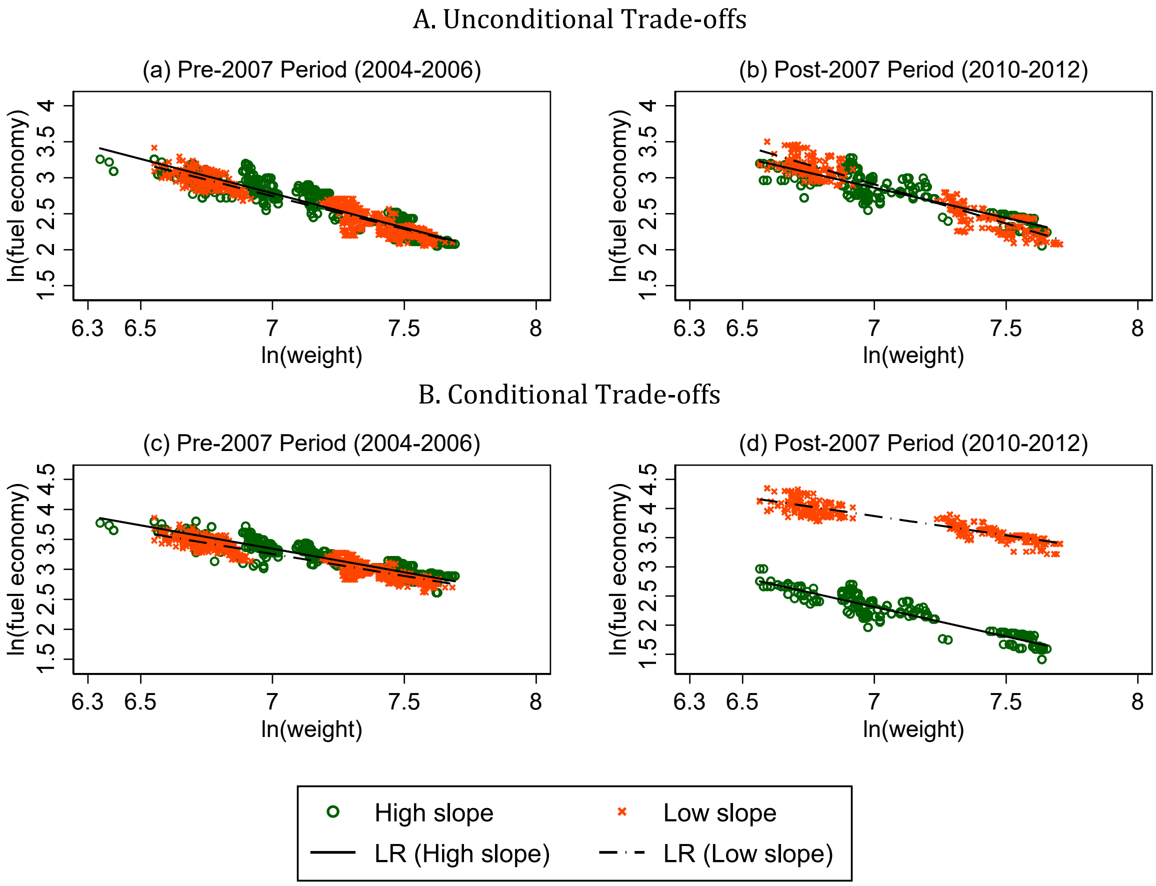
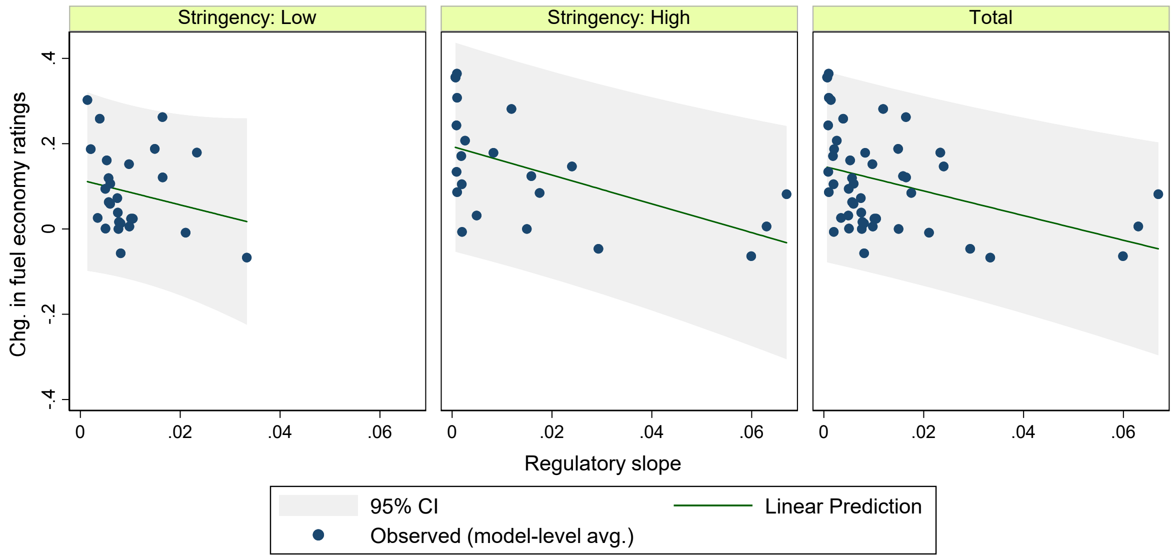


Figure 5. Technical Trade-offs Before and After the New Standards



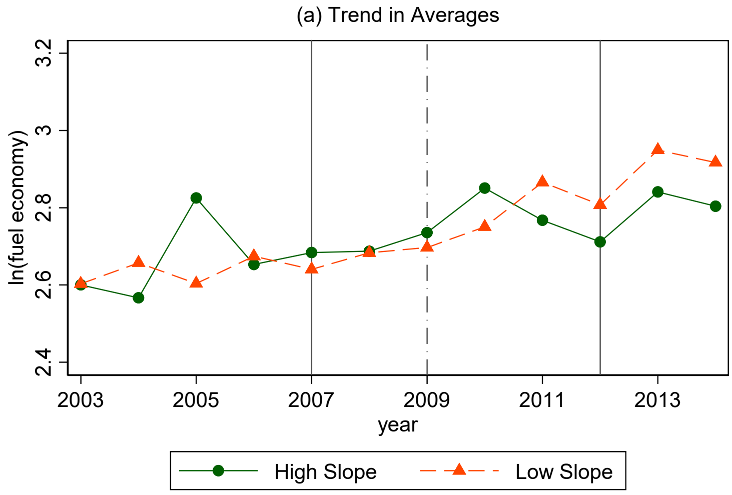
Note: The figure excludes commercial vans and trucks, imported brands, diesel, hybrid, and electric cars as well as observations that fall in weight segments with the first and the third quartiles of compliance costs during the pre-2007 period. Panel A displays scatter plots using raw data. Panel B displays scatter plots using the residuals from a regression of logged fuel economy on key vehicle attributes (in logged values) after removing the linear projection from terms involving horsepower, torque, transmission, and brand dummies.

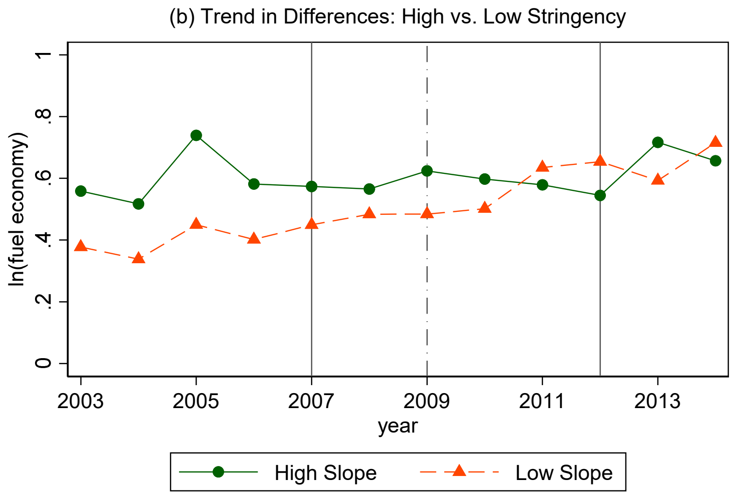
Figure 6. Changes in Fuel-economy Ratings by Regulatory Slope and Stringency



Note: The changes in fuel-economy ratings are calculated as the difference in the model-level means between the pre-2007 and the post-2007 periods. Slope and stringency measures are calculated per equations (8) and (10), respectively.

Figure 7. Trends in Average Fuel Economy between and within Groups





Note: Panel (a) plots average fuel economy ratings in logged values for the high-slope and the low-slope groups. Panel (b) plots the differences in average fuel economy ratings between the high-cost and the low-cost groups for the high-slope and the low-slope groups.

Table 1. Fuel Economy Ratings by Weight Band under the New 2007 Standards

	Weight Segments	H22 (10.15M)	H27 (JC08M)	Slope of Regulation: T	Slope of Regulation: Larger Than Seg. TPF?	Slope of Regulation: T* Mean	Slope of Regulation: T* S.D.	Stringency of Regulation: H	Stringency of Regulation: Quartile of ΔFE	Stringency of Regulation: H* Mean	Stringency of Regulation: H* S.D.	Pre-2007 (2004-06): N	Pre-2007 (2004-06): Fuel Econ. (km/L) Mean	Pre-2007 (2004-06): Fuel Econ. (km/L) S.D.	Post-2007 (2010-12): N	Post-2007 (2010-12): Fuel Econ. (km/L) Mean	Post-2007 (2010-12): Fuel Econ. (km/L) S.D.
1	0 – 600	21.2	22.5	0.12	1	7.95	0.00	1.3	2	1.10	0.00	5	23.40	1.95	0	--	--
2	600 – 702	21.2	21.8	0.00	0	3.99	2.77	0.6	1	1.47	0.47	8	25.88	2.34	0	--	--
3	702 – 740	18.8	21.8	2.11	1	1.74	2.43	3.0	4	1.37	0.50	42	22.12	1.67	43	27.82	4.74
4	740 – 827	18.8	21	0.00	0	3.97	14.18	2.2	2	2.67	1.48	255	20.75	1.94	146	26.12	3.36
5	827 – 856	17.9	21	0.69	1	6.01	18.19	3.1	4	3.35	1.47	150	19.29	1.82	108	23.66	3.41
6	856 – 970	17.9	20.8	0.26	0	4.12	14.97	2.9	4	3.94	1.47	438	18.43	1.70	435	22.80	2.87
7	970 – 1015	17.9	20.5	0.00	0	1.34	1.09	2.6	3	5.10	2.14	130	17.88	3.23	160	20.74	3.49
8	1015 – 1080	16	20.5	2.77	1	1.60	2.36	4.5	4	4.25	1.85	202	17.20	1.70	202	19.37	3.15
9	1080 – 1195	16	18.7	1.30	0	1.71	2.19	2.7	3	3.94	1.56	488	16.35	1.65	318	18.72	2.56
10	1195 – 1265	16	17.2	0.00	0	1.54	1.50	1.2	1	3.42	1.07	310	15.83	1.58	161	17.49	1.99
11	1265 – 1310	13	17.2	3.11	1	1.97	2.68	4.2	4	3.66	1.43	131	13.50	1.82	91	16.20	1.30
12	1310 – 1420	13	15.8	1.27	1	1.89	1.93	2.8	3	3.55	1.44	301	12.97	1.78	162	15.26	1.83
13	1420 – 1515	13	14.4	0.00	0	1.34	1.37	1.4	2	2.98	1.45	339	12.49	1.48	197	13.76	2.29
14	1515 – 1530	10.5	14.4	8.00	1	1.05	0.97	3.9	4	2.99	1.20	75	11.47	1.15	57	12.91	1.06
15	1530 – 1650	10.5	13.2	0.83	0	1.01	0.76	2.7	3	2.58	1.01	419	11.29	1.08	295	12.66	1.58
16	1650 – 1760	10.5	12.2	1.00	0	2.10	5.17	1.7	2	2.58	0.97	281	10.32	1.20	174	12.24	2.34
17	1760 – 1765	10.5	11.1	0.00	0			0.6	1			0	--	--	0	--	--
18	1765 – 1870	8.9	11.1	0.86	1	3.04	7.19	2.2	2	2.07	0.79	193	9.50	0.87	48	10.39	1.34
19	1870 – 1990	8.9	10.2	0.67	0	0.91	0.62	1.3	2	1.60	0.57	141	9.16	0.59	113	10.50	1.19
20	1990 – 2015	8.9	9.4	0.00	0	1.00	0.53	0.5	1	1.72	0.31	16	8.57	0.26	28	10.22	0.85
21	2015 – 2100	7.8	9.4	0.82	1	1.26	0.69	1.6	2	1.65	0.37	32	8.24	0.40	59	9.04	0.55
22	2100 – 2265	7.8	8.7	0.00	0	1.45	0.39	0.9	1	1.59	0.13	14	8.10	0.15	14	8.45	0.61
23	2265 – 2270	6.4	8.7	26.00	1			2.3	3			0	--	--	0	--	--
24	2270 – 3500	6.4	7.4	0.00	0	0.00	0.00	1.0	1	1.19	0.30	11	6.41	0.30	4	7.00	0.12

Table 2. Regression Results using 2004-2012 Passenger Cars

	DD (Pooled)		DD (Stringency: High)		DD (Stringency: Low)		DDD (Pooled)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Bin-level Assignment								
T [Slope: High = 1]	0.035	-0.021	0.036	0.040	-0.001	0.020	-0.002	0.014
	(0.019)	(0.027)	(0.031)	(0.044)	(0.024)	(0.028)	(0.025)	(0.030)
R [Post-2007 = 1]	0.005	-0.047	0.206	0.558	-0.002	0.178	0.004	-0.046
	(0.026)	(0.066)	(0.068)	(0.090)	(0.027)	(0.058)	(0.026)	(0.087)
$R \times T$	-0.035	-0.047	-0.115	-0.239	0.082	-0.002	0.089	0.019
	(0.028)	(0.034)	(0.038)	(0.047)	(0.051)	(0.040)	(0.044)	(0.052)
H [Stringency: High = 1]							0.001	-0.011
							(0.024)	(0.044)
$H \times T$							0.052	-0.043
							(0.037)	(0.045)
$H \times R$							0.088	0.123
							(0.035)	(0.079)
$H \times R \times T$							-0.210	-0.139
							(0.057)	(0.072)
ln(weight)	-1.354	-1.387	-1.429	-1.761	-0.979	-0.748	-1.350	-1.386
	(0.167)	(0.161)	(0.254)	(0.210)	(0.167)	(0.118)	(0.158)	(0.160)
Variant-level controls for observables	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Brand-effect controls	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Segment-effect controls		✓		✓		✓		✓
R^2	0.917	0.930	0.699	0.745	0.933	0.949	0.921	0.932
Obs.	3,253	3,253	1,516	1,516	1,737	1,737	3,253	3,253
Panel B. Model-level Assignment								
T [Slope]	-0.098	-0.083	0.005	0.010	0.153	0.266	0.780	0.433
	(0.019)	(0.019)	(0.018)	(0.018)	(0.049)	(0.040)	(0.134)	(0.076)
R [Post-2007 = 1]	0.012	-0.041	0.162	0.343	-0.099	0.067	-0.043	-0.098
	(0.017)	(0.032)	(0.026)	(0.032)	(0.024)	(0.021)	(0.021)	(0.032)
$R \times T$	-0.200	-0.187	-0.443	-0.402	4.699	2.905	1.414	0.710
	(0.048)	(0.053)	(0.061)	(0.052)	(0.704)	(0.840)	(0.637)	(0.490)
H [Stringency]							-0.021	-0.044
							(0.002)	(0.001)
$H \times T$							-0.184	-0.098
							(0.028)	(0.017)
$H \times R$							0.014	0.012
							(0.005)	(0.005)
$H \times R \times T$							-0.330	-0.187
							(0.131)	(0.101)
ln(weight)	-1.363	-1.324	-1.430	-1.734	-0.946	-0.815	-1.340	-1.032
	(0.043)	(0.052)	(0.071)	(0.069)	(0.055)	(0.062)	(0.040)	(0.034)
Variant-level controls for observables	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Brand-effect controls	✓	✓	✓	✓	✓	✓	✓	✓
Time-varying Segment-effect controls		✓		✓		✓		✓
R^2	0.917	0.930	0.694	0.733	0.933	0.950	0.927	0.951
Obs.	3,247	3,247	1,516	1,516	1,731	1,731	3,247	3,247

Note: Regressions exclude commercial vans and trucks, imported brands, diesel, hybrid, and electric cars as well as observations that fall in weight segments with the first and the third quartiles of compliance costs during the pre-2007 period. In parentheses are clustered standard errors.

Table 3. Regression Results on Vehicle Weight

	DD (Pooled)	DD (Stringency: High)	DD (Stringency: Low)	DDD (Pooled)
	(1)	(2)	(3)	(4)
Panel A: Bin-level Assignment				
DD or DDD	0.002	0.052	-0.004	0.038
Estimate	(0.009)	(0.005)	(0.008)	(0.017)
R^2	0.989	0.975	0.975	0.990
Obs.	2,014	658	1,356	2,014
Panel B. Model-level Assignment				
DD or DDD	0.044	0.075	0.818	0.118
Estimate	(0.025)	(0.026)	(0.352)	(0.061)
R^2	0.988	0.975	0.973	0.988
Obs.	2,008	658	1,350	2,008

Note: All regressions use a subsample consisting only of bins with width less than 40 (in kg). In all regressions, logged curb weight is regressed on the same set of covariates as in Table 2, excluding logged weight. In parentheses are clustered standard errors.

Table 4. Regression Results on Placebo Treatments

	DD (Pooled)	DD (Stringency: High)	DD (Stringency: Low)	DDD (Pooled)
	(1)	(2)	(3)	(4)
Panel A. Perturbing Bin Assignments				
DD or DDD	-0.015	0.014	-0.149	-0.037
Estimate	(0.040)	(0.051)	(0.029)	(0.041)
R^2	0.824	0.728	0.969	0.826
Obs.	2,006	1,130	437	2,006
Panel B. Perturbing Treatment Periods				
DD or DDD	0.014	0.036	0.049	-0.045
Estimate	(0.029)	(0.057)	(0.058)	(0.078)
R^2	0.938	0.737	0.953	0.940
Obs.	2,827	1,183	1,644	2,827

Note: Panel A reports the results of regressions on a placebo treatment where weight cutoffs are shifted by k = 25 kg. Panel B reports on another placebo treatment where the control period is 2003-2004 and the treatment period is 2005-2006. All regressions use the same covariates as in Table 2 including time-varying brand and segment effects. In parentheses are clustered standard errors.

Online Appendix

Appendix A. Proof of Proposition

Proof of (i): We use the following expressions for derivatives, $g_x = \partial g/\partial x$ and $g_{xy} = \partial^2 g/\partial x \partial y$, to economize on space. Let s be given. The first-order condition of the unregulated optimization problem (N) can be rearranged to yield an optimality condition:

$$\frac{U_w}{U_f} = -T_w. \tag{1}$$

That is, the private optimum occurs at the tangency between the indifference curve and the TPF. Given (A1) and (A2), this optimality condition is necessary and sufficient. Because the technology constraint is binding, $f = T(w, s)$ under (A2), the tangency condition above gives us a unique solution to the optimization program given s.

On the other hand, the optimality condition for the social planner's problem (S) is given by:

$$\frac{U_w + \sigma T_w}{U_f} = -T_w. \tag{2}$$

Because $T_w < 0$ and $\sigma > 0$, the socially optimal bundle of attributes equates the slope of the TPF with the slope of the planner's iso-surplus curve that is *flatter* than the firm's iso-surplus curve.

Similarly, the first-order condition of the regulated firm's problem (R) yields:

$$\frac{U_w + \mu\,(T_w - R_w)}{U_f} = -T_w. \tag{3}$$

Comparing (3) and (2), we see that the firm's choice coincides with the social optimum if $\sigma = \mu$ and $R_w = 0$.

Proof of (ii): Let α be the difference between the slopes of R and T [This can be done under assumptions (A1) and (A4)]:

$$\alpha \equiv T_w - R_w. \tag{4}$$

Comparing the three optimality conditions (1), (2), and (3), it is clear that given s, the regulated choice coincides with the unregulated choice if $\alpha = 0$, and the social optimum if $\alpha = T_w$. Note that $R_w \leq 0$, and the value of α ranges from some negative number T_w to some positive number. Furthermore, we see that the regulated choice of attributes occur at the tangency between the TPF and a *flatter* iso-surplus curve than the unregulated iso-surplus curve if $\alpha < 0$ and a *steeper* curve if $\alpha > 0$.

To see the impact of changes in α, totally differentiate (3) with respect to w and α along with $f = T(w, s)$. Rearranging terms, we obtain:

$$\frac{dw}{d\alpha} = -\frac{\mu}{U_{ww} + T_w^2 U_{ff} + 2U_{fw}} > 0.$$

The last inequality follows because the sufficiency of the FOC guarantees that the denominator of the RHS is nonpositive. Hence, given s, the attribute-based regulation tends to bias the attributes of a product to the right (left) of the unregulated attributes if R is steeper (flatter) than T.

Proof of (iii): Next, consider the influence of this bias on the first-stage choice on technology capital s. The logic above suggests that under the regulation, the firm's optimal attribute bundle, in general, depends on both s and α. Let us denote it as $w^*(s, \alpha)$. The regulated firm solves the following maximization problem:

$$\max_{s \geq 0} U(T(w^*, s), w^*) - rs - \mu \left[R(w^*) - T(w^*, s)\right],$$

taking the second-stage solution $w^*(s, \alpha)$ as given. With no regulation, the last term does not show up in the maximand.

By the envelope theorem, we can treat this optimization problem as if w^* is fixed at the optimum. Then the first-order condition yields:

$$(U_f + \mu)T_s = C_s, \tag{5}$$

which simply states that the firm chooses the optimal level of investment, equating the marginal benefit and the marginal cost of investment in fuel-economy technology. It is clear that the condition coincides with the social planner's if $\mu = \sigma$.

To show $ds/d\alpha \leq 0$, differentiating the LHS with respect α, holding s constant, yields:

$$U_{fw} T_s \frac{dw^*}{d\alpha} < 0,$$

where the inequality follows because $U_{fw} < 0, T_s > 0$, and $dw^*/d\alpha > 0$. Hence, the marginal benefit of investment decreases with α, so does the level of investment.

Clarification: A clarifying remark is in order. In the above, we solved the problems as the two-step decision problems. Alternatively, we can formulate them as simultaneous choice problems on (s, w) as follows:

$$\max_{s,w \geq 0} \ U(T(s, w), w) - rs - \mu \left[R(w) - T(w, s)\right].$$

This formulation gives us the same FOC conditions (3) and (5). Substituting (5) into (3), we obtain:

$$\frac{C_s}{T_s} = -\frac{U_w - R_w}{T_w},$$

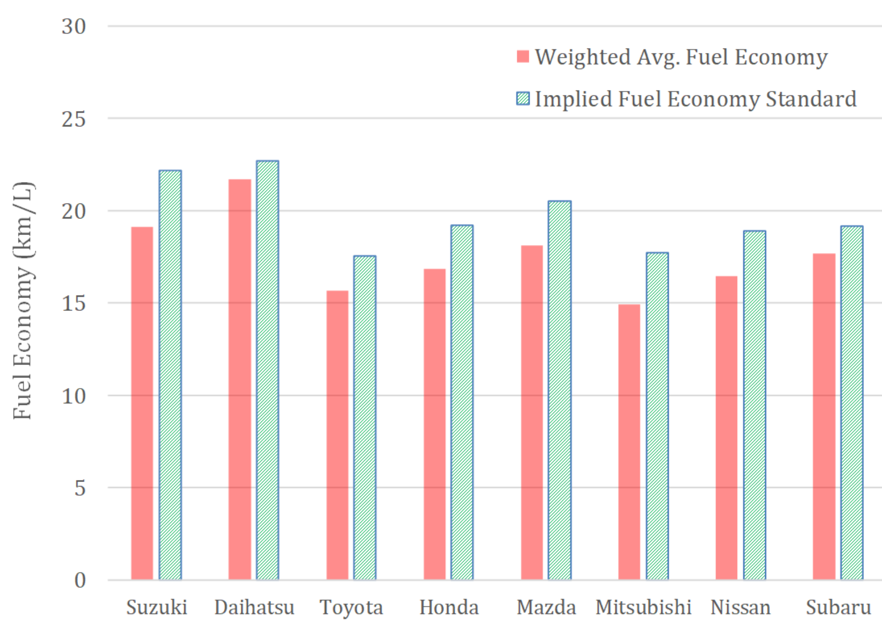
which states that the firm equates the marginal costs of the two compliance strategies, s and w. It is clear from this equation that unless $R_w = 0$, the regulation biases the firm's choice on both s and w. Applying the analogous logic as above, we see that the direction of bias on s depends on the sign of U_{fw}.

Appendix B. Firm-level Compliance as of 2007

We control for firm-level heterogeneity in technical progress by including maker-fixed effects interacted with the time-period dummy in our DDD regression. Yet, if firms differ substantially in compliance level prior to the new standards, they may respond to regulatory treatment quite differently. Since we construct two-fold control-treatment pairs within a narrow weight segment, we do not have sufficient variation to allow for interaction effects to capture this response heterogeneity. We, therefore, check whether firms' compliance levels differed substantially at the beginning of the new standards.

In **Figure A1**, the red bar displays the sales-weighted average fuel economy of vehicles sold in 2007 for each domestic car maker. The green shaded bar reports the estimated fuel economy standard for each maker, using the 2007 sales weights and the 2007 fuel economy standards. These statistics are estimates because we average out fuel economy and weight data over variants of each vehicle model. The exact sales data at the car variant level are not available. The figure demonstrates that at the beginning of the new standards, all domestic car makers were far behind the required fuel economy standards, and hence, are likely to have made some efforts to meet the standards during the post-2007 period.

Figure A1. Sales-weighted Fuel Economy and Standards by Maker in 2007



Appendix C. MLIT Data versus Catalog Data

Section 4 of the main manuscript clarifies differences between our data and the data used in Ito and Sallee (2018). This appendix provides a more detailed account of that discussion. Their data come from the list of new cars published each year by the MLIT. The MLIT list reports data at the car configuration (or '*Katashiki*') level, which is coarser than the grade level reported in the Carsensor catalog. **Figure A2** is a raw image of the original table reported in the Ministry of Land, Infrastructure, and Transportation fuel-economy data. The table captions are in Japanese, so we highlight the relevant section in red. In essence, their data have two major shortcomings: double counting of vehicle offerings and vehicle weights are reported in range for about 3/4 of the car models. The latter is more problematic for our analysis. We thus discuss the extent of this problem in more detail.

Figure A2. MLIT Data Image

ガソリン乗用車(普通・小型)

車名	通称名	型式	原動機 型式	原動機 総排気量 (L)	変速装置の型式及び変速段数	車両重量 (kg)	乗車定員 (名)	10·15モード 燃費値 (km/L)	10·15モード 1km走行における CO2排出量 (g-CO2/km)	10·15モード 燃費基準値 (km/L)	主要燃費改善対策
ホンダ	インサイト	DAA-ZE2	LDA(内燃機関) -MF6(電動機)	1.339	CVT (E)	1190	5	31.0	75	16.0	CY・V・I・EP・C・H
		DAA-ZE2	LDA(内燃機関) -MF6(電動機)	1.339	CVT (E)	1190	5	30.0	77	16.0	CY・V・I・EP・C・H
		DAA-ZE2	LDA(内燃機関) -MF6(電動機)	1.339	CVT (E)	1200	5	28.0	83	16.0	CY・V・I・EP・C・H
	インサイト エクスクルーシブ	DAA-ZE3	LEA(内燃機関) -MF6(電動機)	1.496	CVT (E)	1200～1210	5	26.5	88	16.0	V・I・EP・C・H
		DAA-ZE3	LEA(内燃機関) -MF6(電動機)	1.496	CVT (E)	1210	5	[illegible]	[illegible]	[illegible]	[illegible]
	フィット	DAA-GP1	LDA(内燃機関) -MF6(電動機)	1.339	CVT (E)	1130～1150	5	[illegible]	[illegible]	[illegible]	EP・C・H
		DBA-GE6	L13A	1.339	CVT (E・LTC)	1010	5	24.5	95	17.9	C・V・EP
		DBA-GE6	L13A	1.339	CVT (E・LTC)	1010	5	24.0	97	17.9	C・V・EP
		DBA-GE6	L13A	1.339	CVT (E・LTC)	1030～1080	5	22.0	106	16.0	C・V・EP
		DBA-GE6	L13A	1.339	CVT (E・LTC)	1030～1080	5	21.5	108	16.0	C・V・EP
		DBA-GE6	L13A	1.339	5MT	990	5	21.0	111	17.9	V・EP
		DBA-GE6	L13A	1.339	5MT	990～1010	5	21.0	111	17.9	V・EP
		DBA-GE8	L15A	1.496	CVT (E・LTC)	1070～1090	5	20.0	116	16.0	C・V・EP
		DBA-GE8	L15A	1.496	CVT (E・LTC)	1080～1100	5	20.0	116	16.0	C・V・EP
		DBA-GE8	L15A	1.496	CVT (E・LTC)	1080～1100	5	19.6	118	16.0	C・V・EP
		DBA-GE8	L15A	1.496	CVT (E・LTC)	1080～1100	5	19.2	121	16.0	C・V・EP
		DBA-GE8	L15A	1.496	CVT (E・LTC)	1080～1100	5	18.8	123	16.0	C・V・EP
		DBA-GE8	L15A	1.496	6MT	1050～1080	5	17.4	133	16.0	V・EP
		DBA-GE8	L15A	1.496	5MT	1050～1080	5	17.2	135	16.0	V・EP
		DBA-GE7	L13A	1.339	5AT (E・LTC)	1140～1170	5	17.2	135	16.0	V・EP
		DBA-GE7	L13A	1.339	5AT (E・LTC)	1140～1170	5	17.0	137	16.0	V・EP
		DBA-GE9	L15A	1.496	5AT (E・LTC)	1170～1180	5	16.4	142	16.0	V・EP
		DBA-GE9	L15A	1.496	5AT (E・LTC)	1160～1170	5	16.2	143	16.0	V・EP

Curb weights in kg

Table A1 reports weight distributions in the two data sets. The MLIT data contain a smaller number of observations in each year than our Carsensor data (despite their possible double counting). Of 2,012 observations in the MLIT data between 2010 and 2012, only 25% report exact weights. The remaining 75% of observations report only minimum and maximum weights. The weight range can be as large as 200 kg, averaging at around 35 kg.

When we use observations reported without range, we see, in both data sets, that vehicles are roughly equally distributed to the right and to the left of the 2001 standards' cutoffs, but reported more frequently to the *left* than to the *right* of the 2007 standards' cutoffs. More importantly, when we use observations reported with range (in the MLIT data), *minimum* weights are reported more frequently to the *right* of the 2001 standards' cutoffs for the 2001 standards, yet *maximum* weights are reported more frequently to the *left* of the cutoffs. Interestingly, at the 2007 standards' cutoffs, the frequencies stay the same between minimum and maximum.

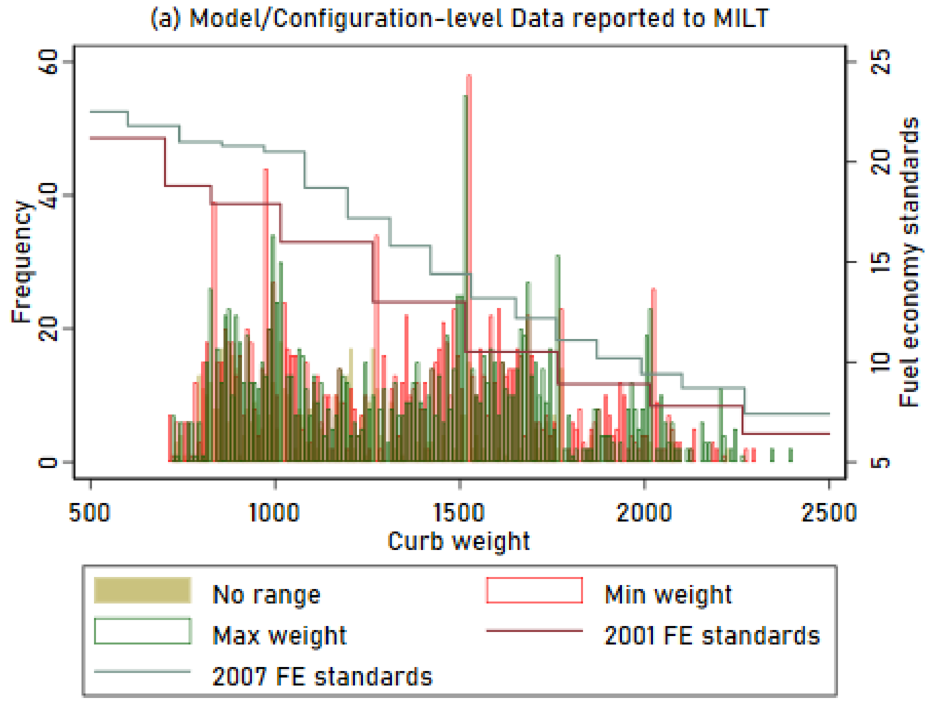
Table A1. Distribution of Vehicle Weight in MLIT Data vs. Catalog Data
2010 – 2012

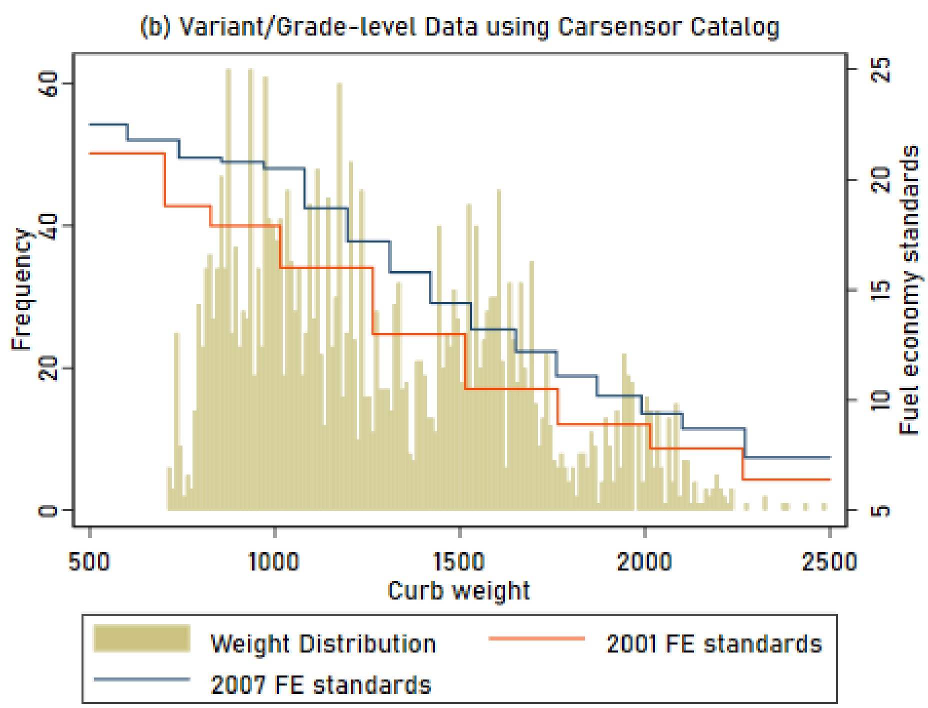
	MLIT Data		Our Data	
	Obs.	Percent	Obs.	Percent
All	2,011		4,303	
Range = 0	508	(0.25)	4,303	(1.00)
Range > 0	1,503	(0.75)	0	(0.00)
Avg. Range (in kg)	35.5			
Min. Range (in kg)	10			
Max. Range (in kg)	200			
Of those reported with range = 0				
Weight at the cutoffs of 2001 standards				
To the right	59	(0.04)	214	(0.05)
To the left	59	(0.04)	192	(0.04)
Weight at the cutoffs of 2007 standards				
To the right	55	(0.04)	439	(0.10)
To the left	159	(0.11)	1,033	(0.24)
Of those reported with range > 0				
Minimum weight at the cutoffs of 2001 standards				
To the right	272	(0.18)		
To the left	59	(0.04)		
Maximum weight at the cutoffs of 2001 standards				
To the right	59	(0.04)		
To the left	235	(0.16)		
Minimum weight at the cutoffs of 2007 standards				
To the right	196	(0.13)		
To the left	555	(0.37)		
Maximum weight at the cutoffs of 2007 standards				
To the right	182	(0.12)		
To the left	560	(0.37)		

Figure A3 visualizes these differences in weight distributions between the two data sets. **Panel (a) of Figure A2** displays three weight distributions using the MLIT data, for all car configurations reported between 2010 and 2012: (i) observations reported with exact weights, (ii) minimum weights (for those with weight ranges), and (iii) maximum weights (for those with weight ranges). **Panel (b) of Figure A2** displays the same using our grade-level data. The figures confirm three points: (1) in the MLIT data, significant bunching occurs at the weight cutoffs, but the incidence of bunching disappears in our data; (2) in the MLIT data, bunching is primarily driven by the observations reported with range, and the minimum weights are clustered at the right of the weight cutoffs while the maximum weights are clustered at the left of the weight cutoffs; and (3) bunching mostly corresponds to the 2001 standards, not the 2007 standards. This last point can be most clearly seen in the weight cutoffs around 1,500 kg. The 2001 weight cutoff around this segment was 1,515 kg whereas the 2007 weight cutoff was 1,530 kg. The bunching is occurring at 1,520 kg, i.e., to the right of the 2001 standard's cutoff and to the left of the 2007 standard's cutoff. We believe the weight distribution in our data is more consistent with findings in the empirical IO literature. For automakers, how best to serve consumer demand and to strategically position and price their products against their market competitors in markets is of the first-order importance. It would not be ideal for automakers to bunch up so many of their vehicles at the weight cutoffs even when they can reduce costs of compliance by doing so.

The question arises naturally then: What explains the behavior in the MLIT data? That is, why do automakers they report the minimum weights so as *not* to cross over to the *lighter* weight category while they report the maximum weights so as *not* to cross over to the *heavier* weight category? Our explanation is as follows. First, the regulatory agency assigns a car model to the lightest weight category when it weights range over two or more weight categories. Therefore, automakers have a very strong incentive not to cross over to the lighter weight category. Second, automakers offer many different variants of the same car model/configuration, yet at the time of reporting the new model data to the MLIT, they do not know how well the new model would perform in the markets, and hence, how many variants of the model they wish to offer, over the course of the model year. Hence, they would like to keep the weight range as large as possible while they would also like to avoid assignment of their models to the lighter weight category. We take these weight distributions as suggesting that the MLIT data offer the evidence of bunching in 'reporting' to the MLIT rather than actual 'product offerings' in the market. Our empirical analysis delivers more convincing evidence on the existence of an incentive to increase weights in actual vehicle offerings. Unfortunately, however, the effect of this incentive is obscured by the other incentive to diversify product offerings, and hence, does not show up as vividly as we wish as bunching at weight cutoffs.

Figure A3. Vehicle Weight Distributions, Years 2010-2012

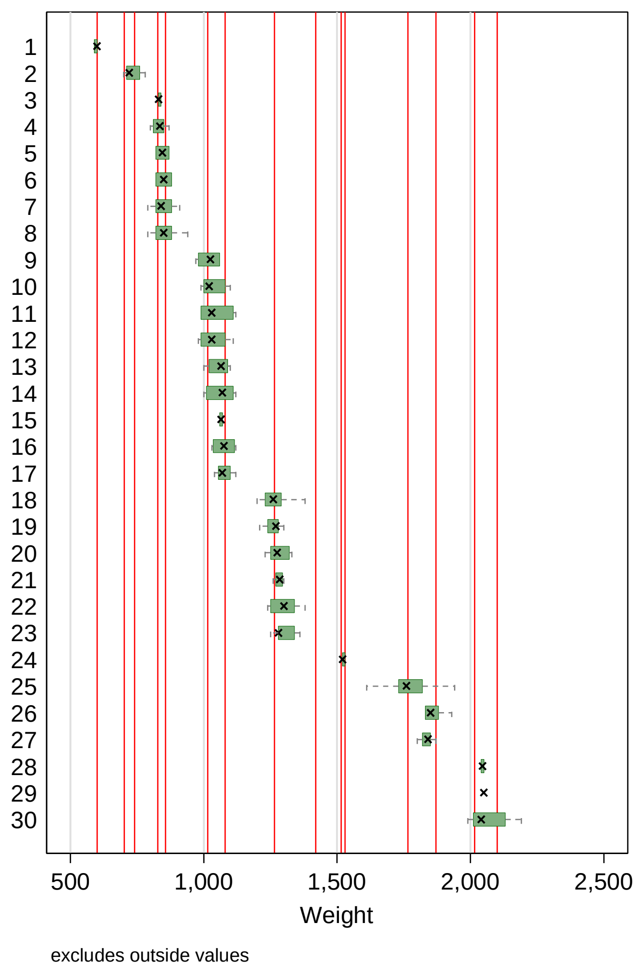




Appendix D. Model-level Assignment

Figure A4 reports the summary of model histories and box diagrams describing the distribution of variant-level curb weights for car models assigned to the high-slope weight bins. Of the 30 models, 11 models did not introduce any new variants between 2010 and 2012, and thus, are classified as 'discontinued'. Of these 11 models, only 2 models had clear successor models. Others either had no clear successor model or were merged to another existing model. The graph demonstrates that for virtually all models, the mean and the median values lie within a single weight bin.

Figure A4. Model History and Distribution of Curb Weights for Vehicle Models Assigned to High Slope Weight Bins

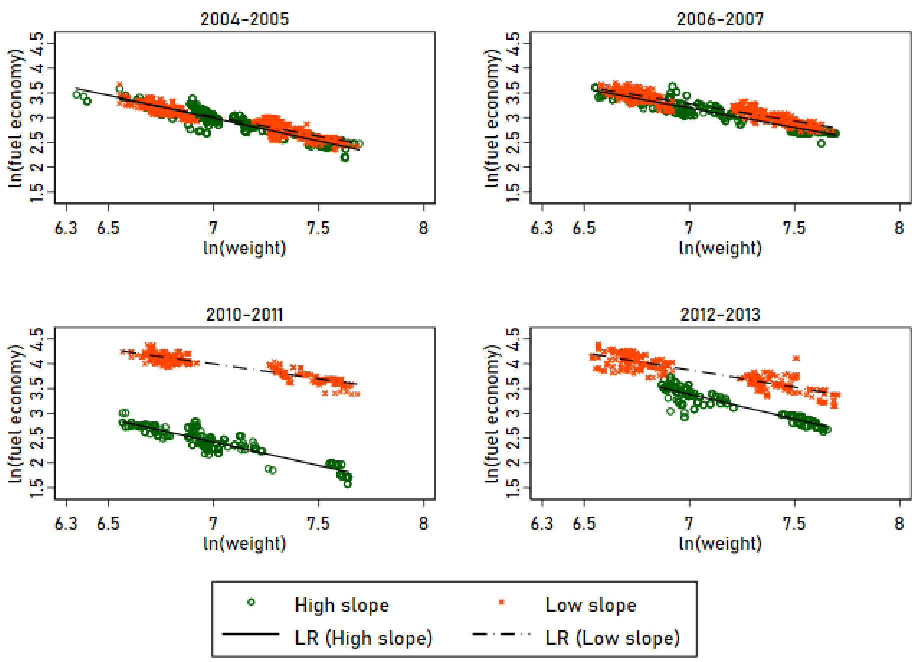


Appendix E. Robustness Analysis

In this appendix, we conduct several robustness checks. The results indeed confirm the robustness of our main results.[1]

Technical Trade-offs: In Figure 5-B of the manuscript, we demonstrate that prior to the new standards, there is no sign of a significant difference in the technical trade-offs, yet after the new standards, the technical trade-offs of those assigned to the high-slope bins lie far below those assigned to the low-slope bins after the new standards. In Figure A5 below, we repeat essentially the same exercise, but using data every two year.

Figure A5. Technical Trade-offs Every Two Year
(Conditional Trade-offs)

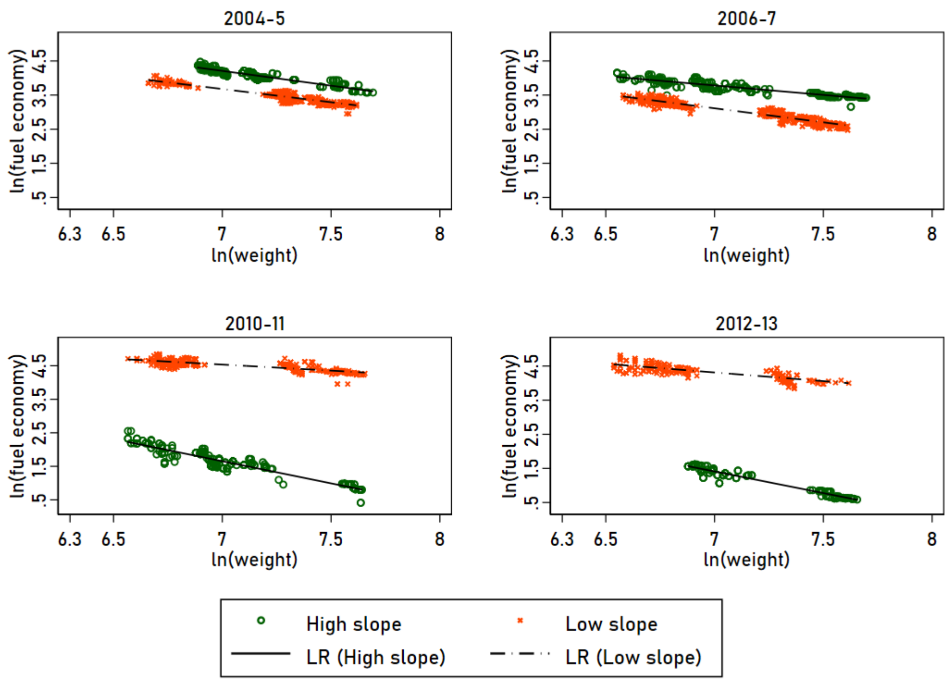


Note: The figure displays scatter plots using the residuals from a regression of logged fuel economy on key vehicle attributes (in logged values) after removing the linear projection from terms involving horsepower, torque, transmission, and brand dummies.

[1] We thank a referee for suggesting these checks.

Sales-weighting: In the manuscript, all regressions are not sales-weighted. Since the enforcement of the regulation is based on firm-level averages, firms may have incentives to respond more for vehicle models with strong demand. Figure A6 replicates Figure A5, using model-level sales as regression weights. The figure seems to confirm our intuition. We, however, do not make use of the sales-weighted regressions in the main analysis because sales data are only reported at the model level with imprecise identifying information, and as a result, we were only able to match the sales data to roughly half of the car attributes observations. Hence, we are still not confident with the exact mechanism for why this occurs. We leave this as a future research agenda.

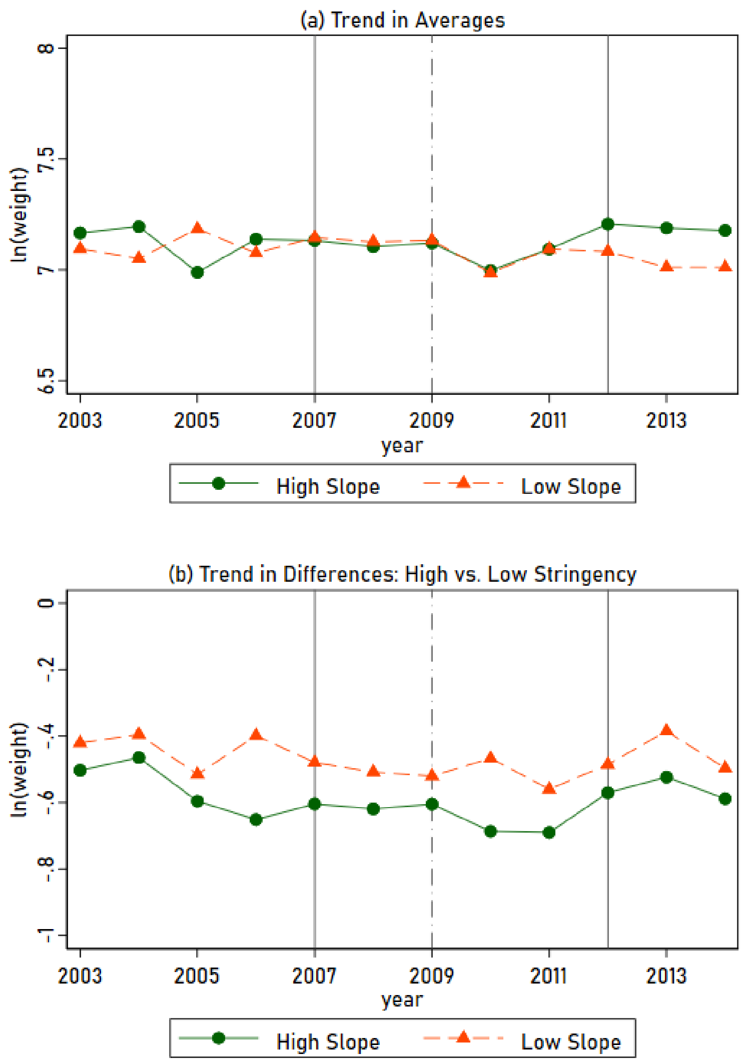
Figure A6. Technical Trade-offs Every Two Year
(Conditional Trade-offs, Sales-Weighted)



Note: The figure displays scatter plots using the residuals from a sales-weighted regression of logged fuel economy on key vehicle attributes (in logged values) after removing the linear projection from terms involving horsepower, torque, transmission, and brand dummies.

Common Trend on Curb Weight: Figure 7 of the manuscript displays the common trends on fuel economy (in log). Figure A7 below repeats the same for curb weight (in log). The top figure indicates that temporal trends (in raw averages) between the two groups are not necessarily identical, yet the differences seem negligible (indeed, smaller than what we observe for fuel economy). The bottom panel shows some bump in 2006 (in differences), but otherwise, trends seem identical during the pre-policy period.

Figure A7. Trends in Average Curb Weight between and within Groups



Note: Panel (a) plots average curb weight in logged values for the high-slope and the low-slope groups. Panel (b) plots the differences in average curb weight between the high-cost and the low-cost groups for the high-slope and the low-slope groups.

'Slope' in Levels versus Logs: In the manuscript, we define our 'slope' variables using raw values for ease of interpretation. However, to be fully consistent with our model, we might define our slope variables using logged values instead. The following table reproduces our main regressions using this alternative definition. It does change the magnitudes of the estimated impacts, yet the results are qualitatively quite similar.

Table A2. Regression on Fuel Economy using Slope Variables Defined in Logged Values

	DD (Pooled)	DD (Stringency: High)	DD (Stringency: Low)	DDD (Pooled)
	(1)	(2)	(3)	(4)
Panel A: Bin-level Assignment				
DD or DDD Estimate	-0.031	-0.239	0.005	-0.113
	(0.029)	(0.047)	(0.027)	(0.067)
R^2	0.929	0.745	0.950	0.931
Obs.	3,391	1,516	1,737	3,391
Panel B. Model-level Assignment				
DD or DDD Estimate	-0.030	-0.056	0.176	-0.291
	(0.007)	(0.007)	(0.047)	(0.176)
R^2	0.930	0.734	0.950	0.953
Obs.	3,247	1,516	1,731	3,247

Note: In parentheses are clustered standard errors.