Comment on the Notice of Proposed Rulemaking for the "Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Years 2021-2026 Passenger Cars and Light Trucks"

Docket No. EPA-HQ-OAR-2018-0283

Docket No. NHTSA-2018-0067

Jeremy J. Michalek
Professor of Engineering and Public Policy
Professor of Mechanical Engineering
Director, Vehicle Electrification Group
Carnegie Mellon University
5000 Forbes Avenue – SH324
Pittsburgh, PA 15213
jmichalek@cmu.edu
(412)268-3765

Kate S. Whitefoot
Assistant Professor of Engineering and Public Policy
Assistant Professor of Mechanical Engineering
Carnegie Mellon University
5000 Forbes Avenue – SH322
Pittsburgh, PA 15213
kwhitefoot@cmu.edu
(412)268-6771

October 26, 2018

Dear DOT Secretary Chao and EPA Administrator Wheeler,

Thank you for the opportunity to provide comments on the proposed rule for passenger car and light truck fuel economy and greenhouse gas emission standards. We are professors at Carnegie Mellon University who have studied vehicle design, economics, environmental impacts, and public policy – including the light duty fleet standards – over the past fifteen years. The views expressed in this comment are provided based on our assessment of the proposed regulations as experts in these areas and are not intended to represent Carnegie Mellon University.

In our assessment, the proposed rule does not satisfy the "maximum feasible" standard required by law, and its analysis of costs and benefits has fundamental flaws that, if resolved, could change agency conclusions about the proposed standards. We detail these concerns below, and we also offer responses to agency requests for comment on several details of the policy. We provide several peer-reviewed scientific publications at the end of this comment that support our assessment.

1. Concerns About the Proposed Rule and the Supporting Analysis

The proposed rule is to freeze fuel economy and greenhouse gas fleet standards at 2020 levels through 2026 instead of allowing them to continue to become more stringent over the period, as defined in current law. We are concerned that the proposed rule does not satisfy the standard set in the Energy Policy Conservation Act. The notice of proposed rulemaking (NPRM) indicates that agency analysis expects the proposed rule to increase petroleum consumption by 0.5 million barrels per day, prevent more than 12,700 fatalities, and reduce driving while making only a small climate change impact and increasing net benefits to society. We are concerned that the analysis has fundamental flaws that, if resolved, could substantially change these estimates and change which policy alternatives maximize net benefits. We will focus on four items:

1.1 Maximum Feasible Standards

NHTSA is required to set the "maximum feasible average fuel economy level" each model year while considering "technological feasibility" and "economic practicability".[1] Although there is ambiguity in determining what level of standard is "maximum feasible", the frozen standard in the proposed rule fails to meet this requirement in a fundamental way because technological capabilities and cost are constantly improving. Capabilities for what is technologically feasible at a particular cost are generally greater in a given year than in years prior. For instance, since 1996, technology improvements have been used to increase fuel economy and/or horsepower of cars by about 2% per year,[2] and the agencies' own preliminary regulatory impact analysis (PIRA) assumes that "manufacturers would still choose to increase fuel economy" every year under the proposed frozen standards.[3] Table 1 shows that the agencies' own analysis predicts that automakers will exceed the standards in every year that the standard remains frozen. Furthermore, automakers are global companies, and they must invest in research and development to meet international regulations on fuel economy and greenhouse gas emissions. Regulations in Canada, China, the E.U., and Japan will continue to increase in stringency in the future, further advancing automakers' technological progress. For all of these reasons, the

---

[1] The Energy Policy Conservation Act of 1975. U.S. Code, Title 49, Subtitle VI, Part C, Chapter 329, Section 32902.

[2] Leard, Linn and Zhou 2017 "The effect of standards for new vehicle fuel economy and GHG emissions on US consumers," *Resources* issue 195, Fall 2017.

[3] Preliminary Regulatory Impact Assessment for The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Year 2021-2026 Passenger Cars and Light Trucks, July 2018, p1016, https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

standard for what is “maximum feasible” increases in stringency over time. The law requires standards to be “maximum feasible,” and frozen standards do not satisfy this criteria. We recommend that a “maximum feasible” standard increase in stringency over time to account for technological advancement and cost reductions.

Table 1: Comparison of agency-estimated CAFE requirements with agency-estimated average fleet fuel efficiency under the proposed standards. The agencies predict that automakers will continue to increase fuel economy every year and will exceed the standards for both cars and trucks in all years that the proposed standards remain frozen (indicated with an *), suggesting that the proposed standards are not “maximum feasible”

	Passenger Cars		**Light Trucks**	
Model Year	Average of OEMs' CAFE requirement from NPRM Table 1-1	Estimated Fuel Economy from PIRA Table 8-34	Average of OEMs' CAFE requirement from NPRM Table 1-2	Estimated Fuel Economy from PIRA Table 8-34
2020	43.7	42.6	31.3	30.7
2021	43.7	43.6	31.3	31.7
2022*	43.7	44.2	31.3	32.0
2023*	43.7	44.5	31.3	32.1
2024*	43.7	44.6	31.3	32.2
2025*	43.7	44.8	31.3	32.3
2026*	43.7	45.1	31.3	32.5

1.2 Fatalities

The agencies claim that the proposed rollback will prevent more than 12,000 fatalities, largely from assumed scrappage of older vehicles due to cheaper new vehicles. Overall, the agencies are assuming that making new cars more expensive leads to *more* cars on the road, but in practice vehicle ownership, driving, and fatalities may actually increase with the proposed rollback.

The scrappage estimate comes from a regression model estimating how new vehicle prices will affect used car scrappage. If tighter standards increase the cost of new vehicles without a sufficient increase in value to consumers (e.g.: due to higher efficiency), this could reduce demand for new vehicles and increase demand for used vehicles, raising used vehicle prices and providing incentives for owners to delay scrappage (but also providing incentives for new and used vehicle buyers to reduce vehicle ownership). The agencies' approach to estimating the magnitude of this potential scrappage effect has several critical flaws that, if resolved, could significantly change the estimated implications of the proposed rule.

First, the regression posed identifies only correlations, not causality. It is likely that other factors not captured by the model, such as changes in employment, economic disparity, or household size, may affect both new vehicle prices and used vehicle scrap rates. If so, then estimating the correlation between new vehicle prices and used vehicle scrap rates does not provide an appropriate model for assessing the effects of a counterfactual scenario in which new vehicle prices are independently increased. We recommend that methods for causal inference be used in counterfactual analysis, or, if causal inference is not possible in this case, that the analysis avoid making causal claims based on non-causal models without adequate emphasis on the potential for bias. One possible direction to reduce potential bias in these estimates is to conduct the regression of used car scrappage on vehicle standards themselves, rather than on new vehicle prices.[4]

Second, the regression has many parameters that are not statistically significant, and the analysis ignores uncertainty. The model, specified in Section 8.10.7.7 of the PIRA estimates the relationship between scrappage rates and vehicle age, new vehicle price, operation cost, and GDP growth, including an assumed functional form with a mix of multiple lag variables, log transformations, and third order polynomial relationships. This specification results in a large number of coefficient

[4] Linn and Dou "How do US passenger vehicle fuel economy standards affect purchases of new and used vehicles?" Report, Resources for the Future, August 2018 http://www.rff.org/files/document/file/RFF%20Rpt%20Linn%20Dou%20Fuel%20Economy%20New%20Cars.pdf

estimates, many of which are not statistically significant. Because the estimated coefficients are uncertain, the effects of changing new vehicle price on used vehicle scrappage rates is also uncertain. Computing scrappage estimates using only point estimates of the coefficients ignores uncertainty, resulting in false precision about the magnitude of the effect. We recommend using a Monte Carlo analysis to understand the distribution of scrappage outcomes implied by uncertainty of the value of the coefficients in the model regression and reporting 95% confidence intervals.

Additional uncertainty stems from model misspecification. For example, by comparing 9000 variations of model specifications, Haaf et al. (2014)[5] show that models of vehicle ownership choices can make substantially different predictions depending on the set and form of the variables used in the assumed model functional form. Unless there is a strong theoretical basis or strong empirical evidence for the particular functional form assumed in the regression, we recommend that the analysis be repeated with multiple alternative plausible functional forms based in the literature to assess how robust the claimed effects are to variation in the assumed model form, and we recommend that the agencies avoid making claims that are not robust to reasonable variation in model specification.

Third, the analysis uses separate assumptions to estimate the effect of annual mileage accumulation and scrappage rates. The agencies request comment on this in the NPRM: "The current model assumes that annual mileage accumulation and scrappage rates are independent of one another. We seek public comment on the appropriateness of this assumption…" The use of independent models effectively assumes that driving patterns are determined by the vehicle rather than by the household.

---

[5] Haaf, C.G., J.J. Michalek, W.R. Morrow, and Y. Liu (2014) "Sensitivity of vehicle market share predictions to discrete choice model specification," *ASME Journal of Mechanical Design* v136 121402 p1-9.

Suppose, for example, that a household owns two vehicles, one new and one old, and decides to scrap the old vehicle early due to lower used vehicle prices induced by relaxing the standards. The agencies' analysis effectively assumes that said household would continue to drive the new vehicle as before, foregoing the travel that used to be provided by the old vehicle, rather than shifting some trips from the old vehicle to the new vehicle. In reality, the household is likely to use the remaining vehicle for at least some of the trips that were previously served by the old vehicle. In assuming that travel patterns are tied to the vehicle, rather than the household, the agencies make a strong assumption that results in implausible predictions, and it serves to overestimate the reduction in vehicle miles traveled and the implications of that reduction, including fatalities. According to internal EPA analysis, for every one extra new vehicle purchased due to reduced costs, the model predicts that 50 additional older vehicles will be scrapped and that the households who lose these vehicles will forego the travel associated with them rather than shift the travel to other vehicles in the household.[6]

We recommend that the agencies either construct and validate an integrated model that accounts for shifts in travel among vehicles with changes in fleet size or, in the absence of such a model, refrain from claiming benefits based on the assumption that travel patterns are tied to vehicles instead of households. We also recommend that the agencies conduct a rigorous, and transparent peer-review of their scrappage assumptions through an independent scientific organization, such as the National Academies.

[6] Preliminary Regulatory Impact Assessment of The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Year 2021-2026 Passenger Cars and Light Trucks, p1016, https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ld_cafe_co2_nhtsa_2127-al76_epa_pria_181016.pdf

1.3 Reduced Driving

It is well known that reducing the cost of driving can induce an increase in driving, and a range of studies over many years have attempted to estimate the magnitude of this "rebound" effect. Prior agency analysis reviewed the literature and used a moderate 10% rebound assumption, but the most recent analysis supporting the proposed rule increases this assumption to 20% based on averaging estimates from studies in the literature from before 2009.[7] The analysis ignores more recent studies that suggest a smaller rebound effect,[8,9,10] it ignores the difference between aggregate rebound and per-vehicle rebound,[11] and it ignores that most studies estimate rebound in response to changes in gasoline prices, whereas rebound in response to changes in vehicle efficiency is likely to be less salient to consumers and result in a smaller effect. The analysis also ignores the effect of changing other costs of driving besides fuel cost – cars that are more expensive also have higher insurance and depreciation costs per mile that affect the cost of driving beyond fuel price. Considering these effects and recent estimates of rebound suggests a smaller rebound effect than assumed in the analysis.

We recommend that the agencies update their rebound assumptions by drawing primarily on recent studies using U.S. data that estimate per-vehicle rebound in response to changes in vehicle efficiency, rather than changes in fuel price.

---

[7] Federal Register Vol. 83, No. 165 p 43100

[8] Gillingham, K., D. Rapson and G. Wagner (2015) "The rebound effect and energy efficiency policy," *Review of Environmental Economics and Policy*, v10 n1 p68-88.

[9] K. Gillingham et al, Nature 493, 475 (2013)

[10] C. Knittel and R. Sandler, American Economic Journal: Economic Policy . Forthcoming

[11] Linn, J., Comment on the NPRM, Docket No. EPA-HQ-OAR-2018-0283 and Docket No. NHTSA-2018-0067

1.4 Technology Projections and Costs

Prior analysis by NHTSA and EPA found that the existing standards increase net benefits to society. But the new analysis ignores EPA's model and uses a changed version of NHTSA's model with many new assumptions that increase the assumed cost of compliance. To our knowledge, the new model has not been independently peer reviewed and assessed for validity of the changes in assumptions.

Agency conclusions that the net benefits of the proposed rule exceed those of the augural standards are based on a number of assumptions about how automakers will implement fuel-saving technologies that are much more pessimistic than prior agency assumptions. In prior analysis it was assumed that automakers would implement the most cost effective fuel saving technologies first, subject to constraints about which technologies are compatible. In the new analysis, internal EPA documents identify several ways that the current model assumes that automakers will implement technologies that are several thousand dollars more expensive than other available packages with similar effectiveness.[12] There are a variety of other pessimistic assumptions in the model, such as the assumption that automakers will convert entire vehicle lines to electric at once rather than offer both electric and gasoline powertrain options, as is common today. Further, the agencies include an assumed loss of value to consumers associated with undesirable attributes of fuel-saving technologies, but a number of fuel saving technologies actually increase performance, and publications in peer-reviewed scientific journals have found that (1) the evidence of hidden costs to vehicle operation characteristics from fuel

[12] SR 12866 Review Materials for The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Years 2021-2026 Passenger Cars and Light Trucks NPRM; RIN 2060-AU09, Office of Management and Budget, https://www.regulations.gov/document?D=EPA-HQ-OAR-2018-0283-0453.

saving technologies is limited[13] and (2) taking advantage of fuel economy / performance tradeoffs while accounting for pricing and consumer demand allows automakers to comply at lower costs than agencies estimate, not higher costs.[14]

We recommend that the agencies conduct a rigorous, and transparent peer-review of their technology cost assumptions through an independent scientific organization, such as the National Academies.

2. Responses to Agency Requests for Comment

2.1 Credit Transfers and Flexibilities

The regulations allow for certain flexibilities, such as transferring "credits" between passenger car and light truck fleets, so that if one fleet over-complies with the regulation, the other can under-comply within a set limit. Automakers can also bank (carry forward) and borrow (carry back) credits within certain time periods. These flexibilities reduce compliance costs with the regulations. However, the agencies' main cost-benefit analysis does not consider these flexibilities over the regulated time period. NHTSA is currently prohibited by statute from including these flexibilities in their standard-setting process.[15] We find this requirement problematic because the automakers use these flexibilities as a

[13] Huang, Helfand, Bolon, Beach, Sha and Smith (2018) "Re-searching for hidden costs: evidence from the adoption of fuel-saving technologies in light-duty vehicles," *Transportation Research Part D: Transport and Environment* v65 p194-212.

[14] Whitefoot, Fowlie and Skerlos (2018) "Compliance by design: influence of acceleration trade-offs on CO2 emissions and costs of fuel economy and greenhouse gas regulations," *Environmental Science & Technology* v51 n18 p10307-10315.

[15] United States. (2007). *Energy independence and security act of 2007*. US Government Printing Office.

common means of complying with the regulation, and ignoring them will bias the cost-benefit analysis to overestimate costs.

2.2 Credit Trading

Automakers who exceed standards earn credits that can be sold to automakers who fall short. The agencies requested comment on whether this credit trading system should be improved or potentially eliminated. Trading provides an opportunity to achieve a given outcome more efficiently – at lower cost, and so we recommend that trading should remain in place and that the agencies should consider options to enable easier trading among automakers. Currently these trades tend to be negotiated one-by-one in an ad-hoc process, which increases transaction costs. A potential route to strengthening efficiency of credit trading is to consider creating a market with transparent prices for trading permits to reduce barriers to trade.

2.3 Driving Cycle Tests and Off-Cycle Credits

The fuel efficiency of each vehicle is measured for regulatory purposes using an old 2-cycle city/highway laboratory test on a dynamometer. These old tests represent a gentle type of driving sensitive to the limits of dynamometers at the time they were developed, and they are not representative of real world driving. The fuel economy ratings consumers see when buying cars are based on 5-cycle tests that account for factors like air conditioning use, aggressive driving, and cold starts. A 50mpg Prius on the new test looks like 70mpg Prius for regulatory compliance purposes, and this creates endless confusion in the public discussion. We recommend that the agencies update the 2-cycle test procedure used for regulatory compliance to match the 5-cycle procedure used to rate vehicles for consumer purchases.

Neither laboratory tests can capture all of the factors that may affect a vehicle's fuel efficiency and emissions. For example, solar reflective glass reduces the load on a vehicle's air conditioner and can save fuel, but this is not captured in laboratory tests. Because the regulation is intended to address real-world fuel consumption and emissions, unless the laboratory test is revised to measure these factors, we recommend that it is appropriate to continue to provide additional credits for technologies that can be shown to reduce fuel consumption and emissions in real world driving. It is important that the credits be applied only to technologies that credibly reduce emissions, so continuing the practice of permitting credits for a pre-approved menu of features that have been tested internally, for technologies demonstrated through A-B testing, and through features with documented evidence of benefit made available for public comment. Better still, we suggest that the agencies consider a redesign of the laboratory test used for compliance using modern data (e.g.: GPS and CAN bus data) to attempt to represent real-world driving as closely as possible.

2.4 Footprint-Based Standards

The current and proposed fuel economy standards set different targets for different vehicles depending on the vehicle's footprint (a measure of size computed as wheelbase × track width). Though designed to encourage adoption of fuel-saving technologies across the fleet without pushing people into smaller vehicles, a peer-reviewed study by Whitefoot and Skerlos[16] suggests they nevertheless encourage automakers to design larger vehicles as a path to compliance, and there is some evidence

---

[16] Whitefoot and Skerlos 2012 "Design incentives to increase vehicle size created from the US footprint-based fuel economy standards," *Energy Policy* v41 p402-411.

that automakers are doing this.[17] We recommend that the agencies reexamine automaker response to the footprint-based standards to determine if adjustments should be made to avoid inducing increases to vehicle size.

A more economically efficient approach of taxing emissions and fuel consumption at socially appropriate levels would allow households to determine whether to reduce fuel consumption and emissions by driving less, by buying a vehicle with more fuel saving technologies, or by buying a smaller vehicle – or, alternatively, to not reduce fuel consumption and emissions at all but rather pay a cost based on the damages they cause. Forcing improvements only through one mechanism (fuel-saving technologies) increases the cost of achieving these outcomes.

2.5 Alternative Fuel Vehicle Incentives

Current standards provide incentives for automakers to produce alternative fuel vehicles, such as electric vehicles, by favorable accounting in compliance calculations. An electric vehicle in 2018 is counted as a zero-emission vehicle in EPA's compliance calculations, ignoring power plant emissions that our research has shown can cause more deaths per vehicle lifetime than emissions from an efficient gasoline vehicle in some parts of the U.S.[18,19] Additionally, one electric vehicle purchase in 2018 is counted as though two electric vehicles were purchased in compliance calculations. Both of these accounting adjustments allow automakers who sell alternative fuel vehicles to have higher

---

[17] United States. Environmental Protection Agency. (2016) Light-Duty Automotive Technology and Fuel Economy Trends: 1975 Through 2016. EPA-420-R-16-010. Page 24.

[18] Weis, A., P. Jaramillo and J.J. Michalek (2016) "Consequential life cycle air emissions externalities for plug-in electric vehicles in the PJM interconnection," *Environmental Research Letters*, v11 n2 024009.

[19] Yuksel, T., M. Tamayao, C. Hendrickson, I. Azevedo and J.J. Michalek (2016) "Effect of regional grid mix, driving patterns and climate on the comparative carbon footprint of electric and gasoline vehicles," *Environmental Research Letters*, v11 n4 044007.

emitting fleets overall and result in increased net emissions every time an alternative fuel vehicle is sold.[20] There are sound reasons to incentivize alternative fuel vehicle development, but these particular incentives do so in a way that leads to higher fleet emissions (even when the alternative fuel vehicle itself is clean), while other approaches, such as tax credits for purchasing electric vehicles, do not necessarily increase emissions. For this reason, we recommend that EPA allow the alternative fuel vehicle incentives to expire and count all sources of life cycle greenhouse gas emissions in compliance calculations with the best available estimates. In contrast, NHTSA handles alternative fuel vehicles differently by statute and counts just 15% of the gasoline-equivalent energy consumed. As electric and other alternative vehicles make up a larger share of the fleet, the distinction between regulating fuel economy and regulating emissions will deepen, and if the government wishes the agencies to continue to harmonize standards, the 15% rule used by NHTSA, which is arbitrary from an emissions perspective, should be revisited.

2.6 Using Net Benefits as a Basis for Policy Selection

The agencies are seeking comment on whether a comparison of net benefits is "an appropriate basis for [policy] selection". Our recommendation is that maximizing net benefits is among the most important factors to consider in policy selection because it is an effort to weigh a variety of policy implications on a common basis and seek decisions that are beneficial to society overall. However, three important caveats should also be considered.

[20] Jenn, A., I.L. Azevedo and J.J. Michalek (2016) "Alternative fuel vehicle adoption increases fleet gasoline consumption and greenhouse gas emissions under United States corporate average fuel economy policy and greenhouse gas emissions standards," *Environmental Science & Technology*, v50 n5 p.2165-2174.

First, estimates of net benefits are inherently uncertain, particularly for policies with such wide-reaching effects as the light-duty vehicle standards, and changes in assumptions can often change the ordering of alternative policy proposals, leaving room for an analyst with an agenda to, intentionally or inadvertently, make assumptions to produce preferred model outcomes. Analysis of net benefits should follow best practices, base assumptions on the peer-reviewed scientific literature, be transparent, clearly justify assumptions, and evaluate a range of scenarios with alternative assumptions to characterize how sensitive estimates are to assumptions. In some cases the uncertainty about net benefits may be too broad to justify policy selection on that basis alone.

Second, while a net benefits analysis captures many aspects of policy impact in a single framework, it does not typically capture all aspects that may be important to policymakers. For example, it does not capture distributional effects of a policy or the potential for the policy itself to induce changes in social norms, preferences, and values.

Third, the legal basis for NHTSA's regulation of vehicle efficiency comes from the Energy Independence and Security Act of 1975, which requires that NHTSA set the "maximum feasible" standards while considering "technological feasibility" and "economic practicability".[21] While net benefits provide a useful metric for policy selection, it is not clear that there is necessarily any relationship between maximizing net benefits and setting the "maximum feasible" criteria while considering "economic practicability".

[21] The Energy Policy Conservation Act of 1975. U.S. Code, Title 49, Subtitle VI, Part C, Chapter 329, Section 32902.

Energy Policy 41 (2012) 402–411

Contents lists available at SciVerse ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Design incentives to increase vehicle size created from the U.S. footprint-based fuel economy standards

Kate S. Whitefoot [a,*], Steven J. Skerlos [a,b]

[a] Design Science, University of Michigan, Ann Arbor, MI, United States
[b] Mechanical Engineering, University of Michigan, Ann Arbor, MI, United States

ARTICLE INFO

Article history:
Received 2 August 2011
Accepted 31 October 2011
Available online 17 November 2011

Keywords:
Fuel economy
Passenger vehicles
Policy incentives

ABSTRACT

The recently amended U.S. Corporate Average Fuel Economy (CAFE) standards determine fuel-economy targets based on the footprint (wheelbase by track width) of vehicles such that larger vehicles have lower fuel-economy targets. This paper considers whether these standards create an incentive for firms to increase vehicle size by presenting an oligopolistic-equilibrium model in which automotive firms can modify vehicle dimensions, implement fuel-saving technology features, and trade off acceleration performance and fuel economy. Wide ranges of scenarios for consumer preferences are considered. Results suggest that the footprint-based CAFE standards create an incentive to increase vehicle size except when consumer preference for vehicle size is near its lower bound and preference for acceleration is near its upper bound. In all other simulations, the sales-weighted average vehicle size increases by 2–32%, undermining gains in fuel economy by 1–4 mpg (0.6–1.7 km/L). Carbon-dioxide emissions from these vehicles are 5–15% higher as a result (4.69×10^{11}–5.17×10^{11} kg for one year of produced vehicles compared to 4.47×10^{11} kg with no size changes), which is equivalent to adding 3–10 coal-fired power plants to the electricity grid each year. Furthermore, results suggest that the incentive is larger for light trucks than for passenger cars, which could increase traffic safety risks.

1. Introduction

In order to reduce the greenhouse gas emissions and oil consumption associated with passenger transportation, the U.S. Congress recently amended fuel economy regulations on new passenger vehicles in the form of the Corporate Average Fuel Economy (CAFE) standards. Responding to criticisms that CAFE encourages the production of smaller vehicles, which unfavorably impacts domestic automakers compared to foreign automakers and may also increase traffic safety risks, the CAFE regulations for vehicles produced from 2011 to 2016 are a function of the footprint (wheelbase by track width) of the vehicles in a manufacturer's fleet such that manufacturers that produce larger vehicles can meet lower fuel economy standards. This regulation design could potentially create an incentive for automotive manufacturers to increase the size of their vehicles and diminish the policy's goal of reduced fuel consumption. Understanding this issue is both important and timely; policymakers are currently developing the CAFE regulations for vehicles produced from 2017 to 2025 and are planning to finalize these regulations by July 2012.

* Corresponding author. National Academy of Engineering, USA. Tel.: +1 202 334 1643.
E-mail addresses: kwhitefoot@nae.edu (K.S. Whitefoot), skerlos@umich.edu (S.J. Skerlos).

0301-4215/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.enpol.2011.10.062

Given these footprint-based standards, a profit-maximizing manufacturer will evaluate various tradeoffs to determine whether modifying vehicle footprint is desirable. These tradeoffs include the marginal reduction in the fuel economy standard, the cost of modifying vehicle footprint, the impact on vehicle fuel economy and other aspects of vehicle performance such as acceleration, and the resulting change in consumer demand. Therefore, any design incentives to modify vehicle footprint will depend on the relationships between these factors.

The National Highway Traffic Safety Administration (NHTSA) states that the dependency of fuel economy targets on vehicle footprint was established such that any incentive to increase or decrease vehicle size would be minimized (NHTSA, 2009). However, despite researchers' recommendations for further investigation (NRC, 2002; Greene and Hopson, 2003), no quantitative analysis was performed to assess what effect the chosen standards have on design incentives to increase or decrease vehicle size. The most closely related analysis examines the impact of weight-based fuel economy standards on changes to vehicle weight (Greene and Hopson, 2003). But, because the relationships between vehicle weight and consumer demand, production costs, fuel economy, and other vehicle attributes are not necessarily the same as the analogous relationships for footprint, their results cannot directly be applied to footprint-based standards.

This study uses simulation analysis to test the hypothesis that the footprint-based CAFE standards will not create an incentive to increase vehicle size. An oligopolistic equilibrium model of the U.S. automotive industry is constructed to study firm incentives in response to the footprint-based CAFE. In this model, firms can adjust vehicle prices, tradeoff acceleration performance with fuel economy, implement fuel-saving technology features, and increase vehicle footprint. The relationships between vehicle performance attributes are determined from engineering vehicle simulations. Results are presented over a wide range of assumptions of consumer preferences for vehicle size, price, acceleration performance, and fuel efficiency.

Changes in the footprint of vehicles have implications for both fuel economy goals and traffic safety. If vehicle footprint increases, gains in fuel economy could be significantly lower. We investigate this issue by determining the change in the sales-weighted average fuel economy observed in simulations that allow firms to increase the footprints of vehicles and comparing this with fuel economy gains assuming, as in NHTSA's (2009) analysis, that vehicle size and sales remain unaffected. With respect to traffic safety, both the absolute measures of vehicle size (the dimensions of the vehicle) and the relative measures of vehicle size (spread of dimensions across vehicles) can impact safety risks (Kahane, 1997; NRC, 2002). This study investigates the impact of footprint-based CAFE standards on both the absolute change in vehicle size and relative differences in vehicle size changes between passenger cars and light trucks, which can be used in conjunction with traffic safety studies to understand the impact of footprint-based CAFE on traffic safety risks.

2. State of the art: CAFE and vehicle footprint or weight incentives

Although researchers have discussed potential design incentives induced by CAFE standards based on vehicle attributes (i.e., vehicle footprint or weight), the majority of these studies are based on qualitative reasoning rather than a quantitative analysis of firm incentives. The National Research Council (NRC) conducted an analysis of CAFE suggesting that the regulations could avoid design incentives to reduce vehicle size or weight by allowing the fuel economy standards to depend on such attributes. Specifically, the analysis reasons that proportionate weight-based fuel economy targets would eliminate motivation for weight reductions, therefore avoiding any adverse safety implications (NRC 2002, see also dissent to this conclusion in Greene and Keller, 2002). The study also mentions that these targets could cause vehicle weight to increase and lead to higher fuel consumption. These conclusions were largely based on regressions of vehicle curb weight on fuel economy and qualitative observations of vehicle weight trends. Additional studies have also raised concerns that attribute-based fuel economy standards could be susceptible to unintended incentives for firms to design vehicles to be larger or heavier in order to qualify for a less stringent standard (Norman, 1994; Greene et al., 2005).

NHTSA constructed the footprint-based CAFE standards using a quantitative analysis but did not study whether manufacturers would have an incentive to change vehicle size as a result of the standards. Fuel economy targets were defined by determining the cost-effective fuel economy that could be obtained without modifying vehicle footprints, and then by fitting a function to these fuel economy values as a function of vehicle footprint (NHTSA, 2006). NHTSA reasoned that, under footprint-based CAFE, if manufacturers redesign a vehicle model to have a smaller footprint, the manufacturer's average fuel economy would increase but so would their required average fuel economy target and, therefore, any incentive to change vehicle footprint would be reduced (NHTSA, 2005, 2006).

Greene and Hopson (2003) analyze the impact of weight-based standards on incentives to increase vehicle weight. In the study, the authors recognize that although manufacturers may be able to lower their required fuel economy standard by increasing vehicle weight, fuel economy also decreases with increased weight. They determine that increasing vehicle weight by 1% would reduce fuel economy performance by 0.6%. Assuming that increasing vehicle weight by 1% would reduce the CAFE requirement by 1% and given a combined standard of 32.7 mpg by 2015, the authors find that the weight-based standard will cause an average increase in weight by 1% and a loss of fuel economy gains by 2.5%.

In addition to studying the footprint-based standards instead of weight-based standards, our approach differs in a few other important ways from Greene and Hopson's analysis. First, we consider the ability of firms to make tradeoffs between fuel economy and acceleration performance and shift production among their vehicle models by modifying prices. This is in addition to changing vehicle footprint and implementing technology features that improve fuel economy at some added cost. Second, we model the automotive industry at a detailed scale, representing all vehicle models and engine options produced in a year by the top twenty firms that sell vehicles in the United States.

3. Methodology

To investigate potential design incentives from the footprint-based CAFE standards, we consider the decisions that an automotive manufacturer may make in response to the regulation. If a manufacturer wishes to increase the footprint of a particular vehicle, the weight of a vehicle will increase to some extent. This will negatively impact both the fuel economy of the vehicle and the acceleration performance. These losses can be alleviated by incorporating various technology features (e.g., lower friction engine components, cylinder deactivation, or lightweight materials) at some additional cost. Another option is to redesign the powertrain to improve fuel economy by compromising acceleration performance, or vice versa. A profit-maximizing manufacturer would balance these decisions based on how the resulting vehicle attributes affect vehicle sales (q), production costs (c), and the ability to meet the CAFE standard. This study is the first analysis of attribute-based standards to consider each of these tradeoffs together.

These decisions can be formulated as an optimization problem where the firm f maximizes profits subject to the constraints of the CAFE regulation. The firm can choose the footprint (ftp), acceleration performance (acc), level of additional technology features ($tech$), and price (p) of each vehicle in their fleet. The constraint of the CAFE regulation is a function of individual vehicle fuel economy targets (T), which depend on the footprint of the vehicle:

$$\max_{ftp_j,acc_j,tech_j,p_j \forall j} \sum_j q_j(p_j,mpg_j,acc_j,tech_j,ftp_j)(p_j - c_j(acc_j,tech_j,ftp_j))$$

subject to

$$\frac{\sum_{j \in \mathfrak{J}_{f,L}} q_j(p_j)}{(\sum_{j \in \mathfrak{J}_{f,L}} q_j(p_j))/mpg_j} \geq \frac{\sum_{j \in \mathfrak{J}_{f,L}} q_j(p_j)}{(\sum_{j \in \mathfrak{J}_{f,L}} q_j(p_j))/T_j} \tag{1}$$

where $mpg_j = f(acc_j,tech_j,ftp_j)$; $T_j = g(ftp_j)$

Because fuel economy, acceleration performance, and the types of technology features incorporated into the vehicle are all related, the above formulation considers fuel economy as dependent on the decision variables acc and $tech$. This choice is arbitrary and equivalent to the manufacturer choosing fuel economy and acceleration performance with the tech variable determined as a function of those attributes.

Demand for a particular vehicle, q_j, in Eq. (1) is dependent upon attributes of the vehicle j as well as the attributes of all

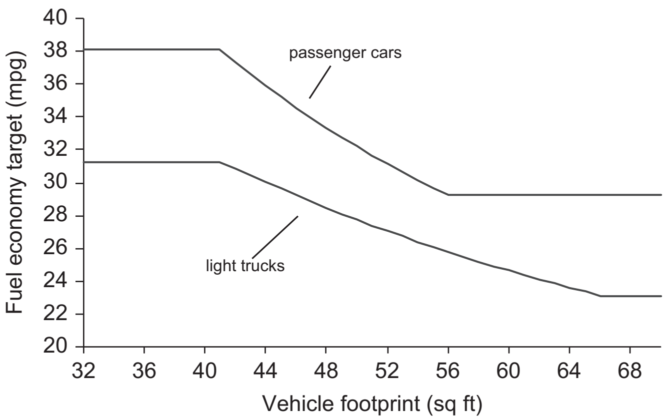


Fig. 1. MY-2014 CAFE footprint-based fuel economy targets.

other vehicles available to consumers. We account for this relationship by solving an oligopolistic equilibrium model where automotive manufacturers seek to maximize profits according to Eq. (1). The subsections below detail how each of the remaining functions in Eq. (1) are derived and how the equilibrium model is formulated.

3.1. Fuel economy targets

The reformed CAFE standards are calculated for each manufacturer as a function of the footprints of the vehicles it produces. Specifically, the regulation sets individual fuel economy targets for each vehicle based on the vehicle's footprint, where larger vehicles have lower targets. A firm will comply with the reformed CAFE standards if the sales-weighted average fuel economy of both its fleets of passenger cars and light trucks are equal to or greater than the respective sales-weighted average targets set for these vehicles as in

$$Stand_L = \frac{\sum_{j \in \Im_{f,L}} q_j}{(\sum_{j \in \Im_{f,L}} q_j)/T_j} \tag{2}$$

The variables $q_j$and T_j in this equation are respectively the sales and fuel economy target for vehicle j in vehicle class L (i.e., passenger cars or light trucks), where the set of vehicles in class L produced by firm f is denoted $\Im_{f,L}$. The model-year (MY) 2014 fuel-economy targets for passenger cars and light trucks as a function of vehicle footprint are described by Eq. (3) and illustrated in Fig. 1:

$$\text{passenger cars}: T_j = 1/min\left(max\left(5.308 \times 10^{-4} \times ftp_j + 4.498 \times 10^{-3}, 1/38.08\right), 1/29.22\right)$$

$$\text{light trucks}: T_j = 1/min\left(max\left(4.546 \times 10^{-4} \times ftp_j + 1.331 \times 10^{-2}, 1/31,30\right), 1/23,09\right) \tag{3}$$

3.2. Tradeoffs between fuel economy, footprint, and acceleration performance

Increasing vehicle footprint leads to a reduction in fuel economy and acceleration performance of the vehicle due to the increase in vehicle weight. We derive these relationships by determining how vehicle weight changes with vehicle footprint and then by determining the relationship between vehicle weight, fuel economy, and 0–60 mph acceleration time. According to Stodolski et al. (1995), approximately 42% of a vehicle's curbweight is attributable to components that are not affected by increases in external vehicle dimensions, such as the engine, transmission, seats, and wheels (also see Kelkar et al., 2001). An additional 9.5% of a vehicle's weight can be approximated as independent of footprint because the height of the vehicle is unaffected.[1] Therefore, a 10% increase in a vehicle's footprint would result in approximately a 5% increase in curbweight. Sensitivity tests of this assumption are described in the Results section.

A regression analysis of the relationship between vehicle footprint and curbweight using MY2006 vehicle data was also performed to compare to this assumption. The estimates of these regression results indicate that, controlling for both engine size and vehicle height, curbweight increases 0.53% with every 1% increase in footprint. Further information on this regression is provided in Appendix A.

The relationship between vehicle weight, fuel efficiency (in gal per 100 mi), and 0–60 mph acceleration time was determined from a combination of physics-based vehicle simulations and data on technology features (e.g., cylinder deactivation). The technology features considered were derived from a subset of technologies identified by NHTSA, which are used to conduct analyses informing the CAFE rulemaking. Table 1 displays a list of the technology features considered in our simulations. The costs of these technology features, estimated by NHTSA (2008), are based on confidential data provided by automotive manufacturers, suppliers, and consultants.

The software package AVL Cruise was used together with these data to simulate the 0–60 mph acceleration time and fuel economy of several vehicle types with varying curbweights, powertrain variables (engine displacement size and the final drive ratio in the transmission), and technology features. Vehicle simulations were conducted for seven separate vehicle segments (i.e. compact cars, minivans, etc.). A total of 29,747 vehicle simulations were conducted to determine how fuel economy and 0–60 mph (0–97 kph) acceleration time change in response to small changes in vehicle curbweight and input powertrain variables. Using these simulation results, the variable tech was created by ordering the cost-effective combinations of technology features that increasingly improve fuel economy and then by assigning an integer value to each ordered combination. This variable represents a continuous approximation of the discrete choices of implementing technology features and is necessary for model tractability. The function **f** in Eq. (1), which describes the relationship between fuel economy, 0–60 acceleration performance, and the level of technology features (tech), was then fit to the simulation data. Further details about this process and the regression can be found in Whitefoot et al. (2011).

Validation tests were performed comparing the approximated relationships based on the simulation data with observed data, shown in Fig. 2. This data includes all non-hybrid vehicle models and engine options in MY-2006. Predicted fuel economy values fit the observed data with an R-squared value of 0.80.

In addition to reducing fuel economy and acceleration performance due to increases in vehicle weight, increasing vehicle footprint may also impact these attributes due to changes in the aerodynamic drag of the vehicle. However, vehicle simulations indicate that a 10% increase in vehicle footprint leads to less than

[1] The body in white, interior less seats, and window glass makes up 35% of vehicle curbweight (Stodolski et al., 1995; Kelkar et al., 2001). We assume that each of these components can be broken down into subcomponents that scale with one side of the vehicle body. Approximating a vehicle as a block with height h, length l, and width w, the surface area of the vehicle body is $2wl+2wh+2lh$. If the footprint increases by 1% the vehicle body's surface area increases by $2.02wl+2\sqrt{1.01}wh+2\sqrt{1.01}lh$. Using model-year 2006 vehicle dimensions, this represents a 0.73% increase in surface area. Therefore, we assume (0.35)(0.27)=9.5% of a vehicle's curbweight depends on the vehicle's height but is independent of the vehicle's footprint.

Table 1
Incremental costs of technology features considered based on NHTSA's (2008) analysis.

Technology costs	Two seater	Compact	Midsize/minivan	Fullsize	SUV	Small pickup	Large pickup/van
Low friction lubricants ($)	3	3	3	3	3	3	3
Engine friction reduction ($)	126	84	126	126	126	126	168
Aggressive shift logic ($)	38	38	38	38	38	38	38
Early torque converter lockup ($)	30	30	30	30	30	30	30
High efficiency alternator ($)	145	145	145	145	145	145	145
Aerodynamic drag reduction ($)	38	38	38	38	38	38	38
Low rolling resistance tires ($)	6	6	6	6	6	6	6
Cylinder deactivation ($)	n/a	n/a	203	203	203	203	229

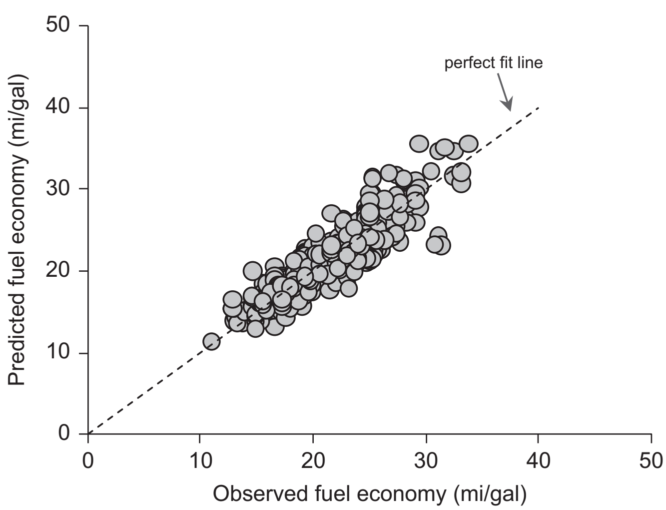


Fig. 2. Comparison of vehicle performance model to MY-2006 vehicle data.

a 1% change in fuel economy and 0–60 mph acceleration time. Therefore, aerodynamic drag is not considered in this study.

The effect of firms using light-weight materials for vehicle components where it is cost-effective under the unreformed CAFE standards are implicitly included in the vehicle performance model. The analysis does not consider the ability of manufacturers to further lightweight their vehicles. NHTSA's (2008) analysis indicate that additional lightweighting is more cost-effective, in terms of fuel-consumption reductions per additional cost, than either engine friction reduction or high-efficiency alternators, but less cost-effective than all other technology features considered. Our results indicate that manufacturers do not implement engine friction reduction or use high-efficiency alternators in the majority of vehicles (60–99% of sales depending on demand parameters), implying that lightweighting is not cost-effective for these vehicles. Therefore, we do not expect the omission of additional lightweighting to substantially affect results. Supposing the contrary, that lightweighting is more cost-effective, then firms would be able to increase the footprint of their vehicles with smaller losses in acceleration performance and fuel economy. Consequently, the extent to which firms have an incentive to increase the size of their vehicles would be, if anything, larger than the results presented in this study. However, the reduction in fuel economy gains may be lower because lightweighting lowers the relative fuel-economy penalty for increasing vehicle size.

3.3. Tradeoffs between footprint and production costs

The product development process for a vehicle model begins with a set of targets specifying vehicle design features, including target vehicle dimensions, followed by detailed design of all vehicle subsystems and ending with vehicle production (Sörenson, 2006; Weber, 2009). The choice of target dimensions at the beginning of this process impacts the resulting production costs of each vehicle in the model line. Most notably, the material costs of the body panels, chassis, glass, driveshaft, axles, and certain interior components will increase with vehicle footprint. Production costs associated with manufacturing processes may also increase. The typical vehicle assembly process involves forming steel sheets into body panels using a series of stamping operations, assembling the panels using robotic arms, spot welding the panels together, and installing subsystem components (Braess and Seiffert, 2005). The costs of these production processes may increase with the vehicle footprint, for example if more time or energy is needed to lift heavier body panels or to provide additional spot welds to assemble the larger panels. Labor costs may also increase if more time is needed to perform assembly operations, for example if additional fasteners are necessary to attach larger subcomponents to the vehicle body.

Acquiring data on these production costs as a function of vehicle footprint is difficult, but we can approximate an upper bound of the impact of increasing vehicle footprint on production costs. Because the aim of this study is to test whether an incentive to increase vehicle size exists, and the extent of this incentive, we use an upper-bound estimate of costs so that our results represent the lower bound of changes to vehicle size. As a conservative upper bound, we assume that increasing vehicle footprint will increase the incremental production costs linearly according to a 1-to-1 relationship, implying that a 1% change in vehicle footprint increases incremental production costs by 1%. We expect that many of the costs of vehicle components and manufacturing operations increase at a smaller rate with vehicle footprint—such as the material costs of body panels—or are completely independent of footprint—such as the costs associated with the seats. Therefore, we expect that this 1-to-1 assumption represents a highly conservative estimate of the impact of vehicle footprint on production costs. If the costs of increasing vehicle footprint are smaller than the assumed relationship, the incentive to increase vehicle size would be larger than results suggest.

Because targets for vehicle dimensions are set early in the product development process and subsequent design of vehicle subsystems considers these dimensions, we do not expect fixed costs associated with vehicle design to increase with incremental decisions on vehicle footprint. We also assume that fixed costs associated with manufacturing processes do not increase with decisions on vehicle footprint. One exception is that the dies used for body-panel stamping scale with footprint dimensions, and therefore the costs associated with the die material increase with footprint. However, the portion of die costs that depend on body panel area is small (Clark and Fujimoto, 1991; McGee, 1973) and so this issue is not considered here.

3.4. Consumer preferences for vehicle size, fuel economy, and acceleration

Consumer demand for new vehicles is modeled as a discrete-choice utility model where consumer utility is a function of vehicle price, fuel consumption, acceleration performance, and vehicle size:

$$U_{nj} = \alpha_1 p_j + \alpha_2 eff_j + \alpha_3 acc_j + \alpha_4 size_j + \xi_j + \epsilon_{nj} \quad (4)$$

Vehicle price, p, in Eq. (4) is measured in ten thousands of 2011 dollars. Fuel efficiency, *eff*, is measured in terms of the gallons of fuel needed to drive 100 miles, and acc is the inverse of the time to accelerate from 0–60 mph (0–97 kph) in tenths of a second, which is approximately proportional to the ratio of horsepower to vehicle weight but also depends on transmission parameters other than horsepower (e.g., the final drive ratio). The parameter size represents the overall length of a vehicle multiplied by the width (L103 by W105 according to SAE International (2005) standards) in ten thousands of sq in. Conversions between footprint and size assume that overall width minus track width, and overall length minus wheelbase, are constant. The ξ_j parameter represents the mean combined utility for all other vehicle attributes, and ϵ_{nj} is an error term specific to individual n and vehicle j.

Multiple confounding factors in observed vehicle and consumer choice data present significant challenges to accurately estimating the α demand parameters. Vehicle prices and observed attributes—including fuel consumption, acceleration performance, and size—are correlated with unobserved vehicle attributes that consumers value, such as exterior and interior styling. This correlation produces biased estimates of the demand parameters. Researchers commonly address this problem by conducting an instrumental variable regression to recover unbiased estimates of the parameters, relying on a set of instruments that are correlated with the observed attributes but are independent of unobserved attributes (e.g., Berry, 1994). However, most of these studies are only concerned with estimating the price parameter; identifying valid instruments for all the attributes listed in Eq. (4) in addition to vehicle prices is particularly challenging (Nevo, 2000). As a result, with only one exception (Klier and Linn, 2008), analyses of CAFE and alternative fuel-economy incentives that estimate consumer preferences have assumed that vehicle attributes other than fuel economy cannot change (e.g., Goldberg, 1998; Jacobsen, 2010; Austin and Dinan, 2005). Therefore, instead of attempting to solve this problem as it would apply to this study, we take a different approach, simulating multiple combinations of values for these preference parameters as scenarios that span the range of reasonably expected consumer preferences as determined by existing literature. While, in many cases, we cannot be certain that these estimates are not biased because of the confounding factors described above, the ranges of estimates in the literature are large enough to presume that they contain the set of plausible values.

Although simulating combinations of demand parameters allows us to investigate the potential incentive to increase vehicle size over multiple scenarios of consumer preferences, this enumeration of demand parameter combinations presents a challenge with regard to computational time. In order to tractably simulate a significant number of combinations of the parameters in Eq. (4), it is necessary to make a simplifying assumption that the α coefficients are common across all consumers, meaning that heterogeneous preferences are not accounted for in this model. Following customary assumptions of the logit model, the ϵ_{nj} parameters are assumed independently and identically distributed across vehicles according to a Type 1 extreme value distribution. This assumption allows the expected value of sales of vehicle j to be written as in

$$\mathbb{E}(s_j) = N \frac{e^{V_j}}{e^{V_{og}} + \sum_{k \in \Im} e^{V_k}} \quad (5)$$

$$V_j = \alpha_1 p_j + \alpha_2 gpm_j + \alpha_3 acc_j + \alpha_4 size_j + \xi_j$$

The parameter N in Eq. (5) is the number of consumers, $\Im$ is the set of vehicles in the market including vehicle j, and V_{og} is the utility of the outside good, representing the utility of not purchasing a new vehicle. Given the sales of vehicle $j(s_j)$; the number of consumers that did not purchase a new vehicle (s_{og}); and values of the α coefficients for price, fuel consumption, acceleration performance, and size, the mean utility of all other vehicle attributes (ξ_j) can be inferred as

$$\xi_j = \log\left(\frac{s_j}{I}\right) - \log\left(\frac{s_{og}}{I}\right) - (\alpha_1 p_j + \alpha_2 gpm_j + \alpha_3 acc_j + \alpha_4 size_j) \quad (6)$$

The ranges of plausible values for the α coefficients in the equations above were determined based on key properties of consumer demand for new automobiles estimated in the literature. Ranges for the price coefficient were based on estimated values for the average price-elasticity of demand, which range from -2.0 to -3.1 in the literature (Berry et al., 1995; Goldberg, 1998; Jacobsen, 2010; Klier and Linn, 2008; Train and Winston, 2007). Ranges of values for the remaining coefficients were informed based on the willingness of consumers to pay for improved fuel consumption, faster acceleration performance, and larger size as estimated from the literature. These estimates were either derived from logit models that consider consumer preferences to be homogeneous or random-coefficient logit models where the mean of the distribution is used to derive willingness-to-pay.[2] The average estimated willingness to pay for vehicle attributes ranges from \$340 to \$2000 for an additional sq ft of vehicle size (\$366–\$2150 per 1000 cm^2), \$160–\$5500 for an increase of 0.01 hp/lb in acceleration performance (\$97–\$3345 per 0.01 kW/kg), and \$1100–\$9000 for a reduction in fuel consumption of 1 gal per 100 miles (\$468–\$3826 per L/100 km) (Beresteanu and Li, 2008; Greene and Liu, 1987; Klier and Linn, 2008).

Helfand and Wolverton (2009) recently conducted a survey of consumer valuation for fuel economy and found that estimates for consumers willingness to pay for 1 mpg (0.43 km/L) more of fuel economy ranges from approximately \$200–\$600 in the literature. Using the vehicle data input into our simulations, this corresponds to an average willingness to pay as low as \$800 for improved fuel efficiency of 1 fewer gal per 100 miles (\$340 per L/100 km), which is less than the lower bound determined above. Therefore, we use \$800-\$9000 for 1 fewer gal per 100 miles (\$340–\$3826 per L/100 km) as the range of consumer preference for fuel efficiency instead.

Table 2 reports the ranges of willingness-to-pay for vehicle attributes as estimated in the literature and the α coefficients that correspond to these ranges. Ideally, combinations of these parameters for the simulations would be determined by sampling from their joint distribution. However, existing literature has neither produced estimates of this joint distribution nor characterized correlations between these parameters. Consequently, combinations of these parameters were simulated assuming independence of preference parameters so as to span the complete range of consumer preference scenarios that would be produced using any correlation of parameters. Specifically, the parameter ranges were divided up into three levels for each parameter—representing the lower bound, midpoint, and upper bound for each parameter—and combinations of these parameter levels were used as simulation inputs. Assuming that the incentive to change vehicle size is monotonic with consumer preferences for vehicle size, price, fuel efficiency, and acceleration performance, the range of the results of this study bound the

[2] Boyd and Mellman (1980), referenced in Greene and Liu's (1987) review, estimate a random-coefficient utility model using a lognormal distribution on the coefficient of fuel efficiency. In this case, the median of the distribution was used to derive willingness-to-pay values.

Table 2
Ranges of demand parameters in literature and corresponding model coefficients.

	Range of mean price elasticity		
	Low	Mid	High
Price	2.0	2.6	3.1
Corresponding coefficients, α_{price}	0.65	0.83	1.00
	Range of estimated willingness to pay		
	Low	Mid	High
Vehicle Size (sq ft)	\$340	\$1170	\$2000
(sq m)	\$366	\$1259	\$2153
Corresponding coefficient			
$\alpha_{price}=1.00$	2.12	7.44	12.71
$\alpha_{price}=0.83$	1.76	6.18	10.55
$\alpha_{price}=0.65$	1.38	4.84	8.26
Acceleration performance (0.01 hp/lb)	\$160	\$2830	\$5500
(0.01 kW/kg)	\$97	\$1721	\$3345
Corresponding coefficient			
$\alpha_{price}=1.00$	0.06	1.07	2.07
$\alpha_{price}=0.83$	0.05	0.89	1.72
$\alpha_{price}=0.65$	0.04	0.70	1.35
Fuel efficiency (gal/100 mi)	\$800	\$4900	\$9000
(L/100 km)	\$340	\$2083	\$3826
Corresponding coefficient			
$\alpha_{price}=1.00$	0.07	0.44	0.80
$\alpha_{price}=0.83$	0.06	0.37	0.66
$\alpha_{price}=0.65$	0.05	0.29	0.52

results that would be produced using any combination of demand parameters in the ranges specified. Strong evidence supporting this monotonicity is shown in the Results section.

3.5. Equilibrium model

Producer decisions regarding vehicle prices and attributes are modeled as an oligopolistic equilibrium model where firms maximize profits with respect to the prices, acceleration performance, and levels of technology features of their vehicles. The top twenty automotive firms that sell vehicles in the United States are represented in the model. Vehicles are represented as all vehicle models and engine options produced by these firms based on MY-2006 data, totaling 473 vehicles.

Firms are differentiated as to whether they are expected to meet the CAFE standards even if it is more profitable to violate them. The model allows BMW, Jaguar, Mercedes-Benz, Porsche, and VW to violate the standard and pay the legally required penalties. The profit maximization formulation for these firms takes the form of

$$\max_{ftp_j, acc_j, tech_j, p_j \forall j} \sum_j q_j(p_j - c_j) - F_C - F_T \tag{7}$$

where

$$mpg_j = f(acc_j, tech_j, ftp_j)$$

$$F_C = \left(\sum_{m \in \mathfrak{C}_C} q_m\right)\left(\sum_{m \in \mathfrak{C}_C} \frac{q_m}{q_m/T_m} - \sum_{m \in \mathfrak{C}_C} \frac{q_m}{q_m/mpg_m}\right)$$

$$F_T = \left(\sum_{n \in \mathfrak{C}_T} q_n\right)\left(\sum_{n \in \mathfrak{C}_T} \frac{q_n}{q_n/T_n} - \sum_{n \in \mathfrak{C}_T} \frac{q_n}{q_n/mpg_n}\right)$$

The parameters F_C and F_T are, respectively, the penalties for violating the fuel economy standard for passenger cars ($Stand_C$) and light trucks ($Stand_T$). Fuel economy targets, T_m and T_n, for these vehicle classes are determined by Eq. (2). All other firms are treated as constrained to the CAFE standards so that their profit maximization problems take the form of Eq. (3).

Firm decisions on vehicle footprint are constrained to a maximum of a 10% increase. This constraint is imposed to avoid extrapolation outside of the boundaries of the data used to construct the engineering performance model and to account for any potential manufacturing constraints of dramatically increasing vehicle footprint. Data of vehicle models from 1997–2010 indicate that increases in vehicle footprint by 10% compared to the previous model design occur (Chrome Systems, Inc., 2008), suggesting that any potential constraints on footprint are at least 10% and, therefore, imposing this constraint on the model causes the results to represent a lower bound with respect to the incentive to increase vehicle size under the footprint-based CAFE standards.

4. Results

Simulations were performed for a number of combinations of consumer preference parameters and the change in the sales-weighted average of overall vehicle size (length by width) across all vehicle models was determined. Table 3 presents results for scenarios in which the average price-elasticity of demand is high. This represents a conservative case in which incentives to increase vehicle size are lower because consumers are not as willing to pay for the cost of increasing vehicle size. The upper left corner of this table represents the lower bound of changes in vehicle size caused by the MY-2014 footprint-based CAFE standards. The table also illustrates how changes in vehicle size vary with different levels of consumer preferences for vehicle size, fuel efficiency, and acceleration performance.

For the results in Table 3, consumer preference parameters for acceleration performance and fuel efficiency are set at the same level (i.e., either both low, both high, or both at midpoints). Additional simulation results are presented in Table 4. These results illustrate the sensitivity of changes in vehicle size under footprint-based CAFE to consumer preference parameters, including independent variations of preference for fuel efficiency and acceleration performance. The last

Table 3
Changes in sales-weighted average vehicle size given combinations of consumer preference parameters with price sensitivity at the upper bound.

				Preference for vehicle size		
				Low	Mid	High
Preference for fuel efficiency	High	Preference for acceleration	High	−1.4 sq ft (−0.13 sq m)	+3.8 sq ft (+0.35 sq m)	+7.0 sq ft (+0.65 sq m)
	Mid		Mid	+1.5 sq ft (+0.14 sq m)	+7.5 sq ft (−0.70 sq m)	+9.2 sq ft (+0.85 sq m)
	Low		Low	+2.1 sq ft (+0.20 sq m)	+9.6 sq ft (+0.89 sq m)	+13.4 sq ft (+1.24 sq m)

Table 4
Sensitivity of results to variations in consumer preference parameters.

Price sensitivity	**Preference for fuel efficiency**	**Preference for acceleration**	**Preference for vehicle size**	**Sales-weighted average change in footprint**	
High	Mid	High	Mid	+4.0 sq ft	(+0.37 sq m)
High	Mid	Low	Mid	+9.4 sq ft	(+0.87 sq m)
High	High	Mid	Mid	+5.9 sq ft	(+0.55 sq m)
High	Low	Mid	Mid	+9.2 sq ft	(+0.85 sq m)
Mid	Mid	Mid	Mid	+10.5 sq ft	(+0.98 sq m)
Low	Mid	Mid	Mid	+11.3 sq ft	(+1.05 sq m)
High	Low	High	Mid	+5.9 sq ft	(+0.55 sq m)
High	High	Low	Mid	+9.3 sq ft	(+0.86 sq m)
High	Mid	High	Low	−1.0 sq ft	(−0.09 sq m)
High	High	Mid	Low	+1.3 sq ft	(+0.12 sq m)
Mid	Mid	Mid	Low	+4.2 sq ft	(+0.39 sq m)
Low	Low	Low	High	+16.1 sq ft	(+1.50 sq m)

line of Table 4 represents the upper bound of changes in vehicle size caused by the MY-2014 footprint-based CAFE standards.

Results indicate that there is an incentive to increase vehicle size in all simulations except the scenarios in which consumer preference for size is at the lower bound ($340 per sq ft) and preference for acceleration performance is at the upper bound ($5500 per 0.01 hp/lb). In those cases, firms have an incentive to shift production of their vehicles such that the average vehicle size decreases by 1.0–1.4 sq ft (0.09–0.13 sq m) due to low consumer preference for vehicle size compared to acceleration performance. In all other simulations, firms have an incentive to increase the size of vehicles sold, both by increasing the footprint of vehicle models and by shifting production toward larger vehicles. The incentive varies substantially depending on consumer preferences, from an average of 1.4–16.1 sq ft (0.13–1.21 sq m). This compares with an average increase in size of 1 sq ft (0.09 sq m) between 2008 and 2011.

Depending on the scenario, between 7% and 33% of vehicle models and engine options are actively constrained by the 10% upper bound on the increase in vehicle footprint. The larger percentage occurs in scenarios where consumer preference for vehicle size is high and preferences for acceleration performance and fuel efficiency are low. This suggests that the increase in vehicle size for these scenarios would be even higher if this constraint was relaxed.

Sensitivity of these results with respect to assumptions on vehicle weight and production costs was also investigated. Vehicle curbweight was assumed to increase by 0.5% for every 1% increase in footprint. A 40% variation in the percentage of vehicle weight that changes with footprint leads to less than a 5% change in results. Production costs were assumed to increase 1% for every 1% increase in footprint as a highly conservative upper bound. Assuming instead that an increase in footprint by 1% increases production costs by 0.8%, the change in average vehicle size is approximately 9% greater.

To test the impact of the incentive to increase vehicle size on fuel economy, we compare simulation results to the average fuel economy that the CAFE standards would require if vehicle size and sales remain unaffected. Specifically, the sales and vehicle footprint using MY-2006 data was input into Eqs. (1) and (2) to determine these fuel economy standards. This is similar to the process NHTSA has used to predict future levels of fuel economy, except they have used product development plans provided by automotive firms to extrapolate future vehicle attributes. Our calculations from this procedure indicate that the required average fuel economy under the MY-2014 footprint-based standards is 30.7 mpg (13.1 km/L). This is similar to NHTSA's estimated value of 31.5 mpg (13.4 km/L). Simulation results indicate that the combination of increases in vehicle size and shifts in production to larger vehicles can reduce these fuel economy requirements. The resulting required fuel economy standards from the simulations are 1.4–4.1 mpg (0.6–1.7 km/L) lower than if vehicle sales and size remained unaffected.

Simulations results also suggest that the incentive to increase vehicle size is significantly different for light trucks and for passenger cars. Fig. 3 illustrates the change in vehicle footprint and fuel economy from simulation results using midpoint values of consumer preference for fuel efficiency, acceleration performance, and vehicle size. Initial vehicle data is displayed in gray with counterfactual simulation results in black. The sizes of the circles in the figure are proportional with vehicle sales. The sales-weighted harmonic mean of fuel economy and vehicle footprint are plotted as a cross (+).

The figure illustrates that vehicle footprint increases for both passenger cars and light trucks, but that the increase in footprint

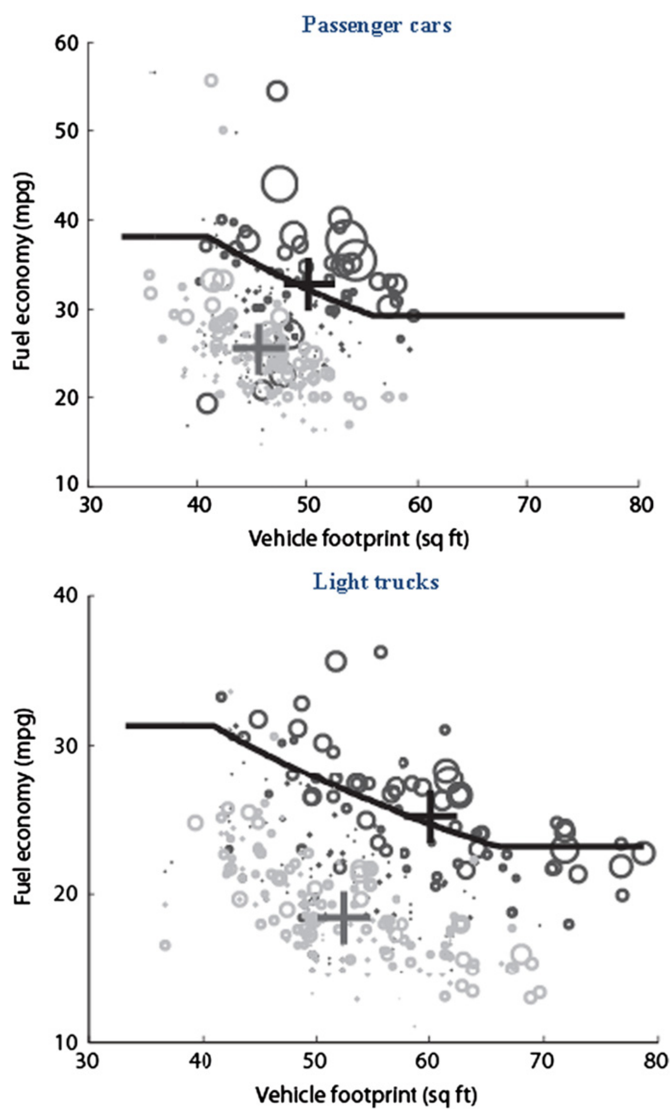


Fig. 3. Simulation results given midpoint consumer preferences. Sales-weighted harmonic mean vehicle attributes are represented as a cross (+). Initial data are in light gray, with MY-2014 CAFE counterfactual simulation results in dark gray. Circle size is proportional to vehicle sales.

for light trucks is significantly larger than for passenger cars. The sales-weighted average increase in vehicle footprint is 9.9 sq ft (0.92 sq m) for light trucks but 5.7 sq ft (0.53 sq m) for passenger cars.

This behavior can be explained by the larger impact of the CAFE standard for light trucks on firm profits than the standard for passenger cars. Simulation results give the Lagrange multiplier to the constraints in eq. (3), which is interpreted as the incremental profit loss given an incremental increase in the CAFE standard, referred to as the shadow cost of the standard. Results indicate that this shadow cost is 1.5–7.0 times larger for light trucks than passenger cars. Because the light truck standard causes larger profit losses than the passenger car standard, firms increase the sales-weighted average footprint of light trucks more than passenger cars in 20 out of the 21 simulations conducted.

Similar counterfactual simulations for the reformed CAFE standards have not been performed; so these shadow costs cannot be compared to other estimates in the literature. With regard to the unreformed CAFE standards, Anderson and Sallee (2009) also found that the ranges of estimated shadow costs of the standard for light trucks were larger than for passenger cars for Ford, GM, and Chrysler. Jacobsen (2010) found that the shadow cost for light trucks was larger than passenger cars for Ford, but that the shadow cost for light trucks was lower than for passenger cars for GM and Chrysler.

The incentive to increase vehicle size also varies substantially among vehicle models within the same class. For the case illustrated in Fig. 3, in which consumer preferences for vehicle size, fuel efficiency, and acceleration performance are all at their mid-points and price-elasticity for demand is high, increases in vehicle footprint range up to 13.8 sq ft (1.28 sq m) for certain light-truck models, and 10.4 sq ft (0.97 sq m) for certain passenger-car models. Even in the cases in which the sales-weighted average vehicle size decreases, the size of certain vehicle models increase by as much as 8.5 sq ft (0.79 sq m).

Additional simulations were performed to test the impact of changing the slope of the functions determining fuel-economy targets dependent on vehicle footprint, as described by eq. (3). These functions were iteratively modified to decrease the slopes until simulation results show no increase in the sales-weighted average footprint for the case where consumer preferences for vehicle size, fuel efficiency, and acceleration performance are all at their midpoints. Results indicate that if the slope of the function for passenger cars is reduced by a third and the slope of the function for light trucks is reduced by half, then the sales-weighted average footprint does not increase for this scenario of consumer preferences. Fig. 4 illustrates these results.

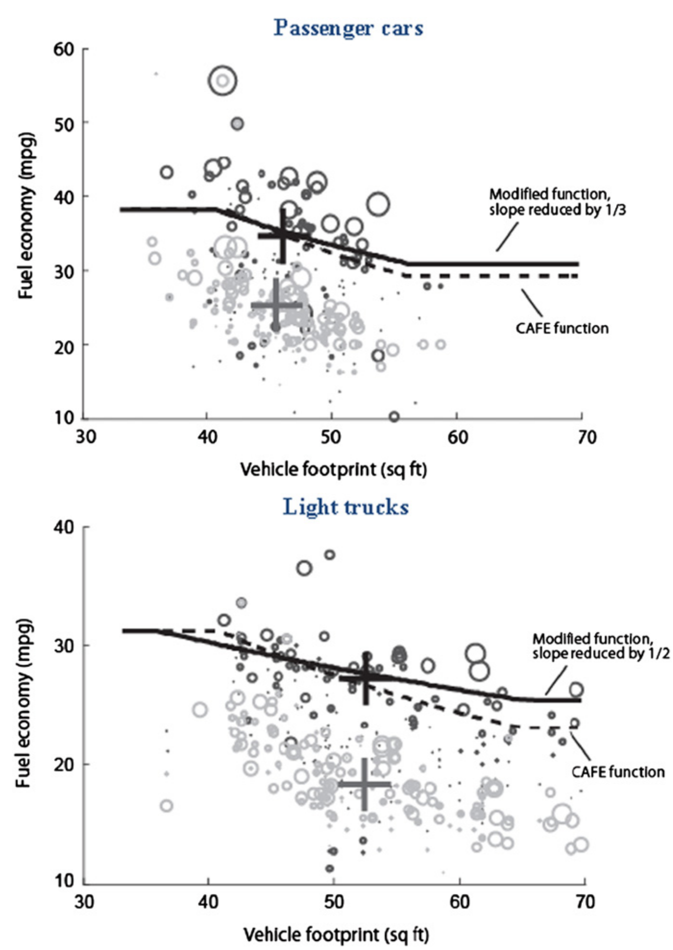


Fig. 4. Simulation results for midpoint consumer preferences with modified functions determining fuel-economy targets dependent on vehicle footprint. Sales-weighted harmonic mean vehicle attributes are represented as a cross (+). Initial data are in light gray, with MY-2014 CAFE counterfactual simulation results in dark gray. Circle size is proportional to vehicle sales.

5. Discussion

This analysis shows that the current footprint-based CAFE standards create an incentive to increase vehicle size that undermines gains in fuel economy over a large range of assumptions about consumer preferences. The hypothesis that the footprint-based CAFE standards do not create an incentive to increase vehicle size can be rejected except under somewhat extreme simultaneous assumptions regarding consumer preferences for vehicle size and acceleration performance. Assuming vehicles are driven 12,000 miles per year for 10 years and annual U.S. new vehicle sales are 13 million, results indicate that the reduction in required fuel economy caused by the incentive to increase vehicle size leads to an additional 24–76 million short tons (22–69 Mtonnes) of annual CO_2 emissions—comparable to adding 3–10 coal-fired power plants (each 1000 MW) to the electricity grid each year (Fay and Golomb, 2002).[3]

The results also suggest that the incentive to increase vehicle size is greater for light trucks than for passenger cars, which would increase the divergence of the sizes of vehicles in these classes. This divergence could negatively affect traffic safety because one can expect a divergence in the weight of vehicles in these classes corresponding to their divergence of size. Although the literature on traffic safety has not produced a consensus on the relationship between vehicle size and safety, researchers generally agree that if the spread of vehicle weight on the road increases, fatality risk in a two-vehicle crash increases (Anderson and Auffhammer, 2011; Greene and Keller, 2002; Kahane, 1997).

While the footprint-based CAFE standards can theoretically be modified to eliminate incentives to change vehicle size, this study illustrates that this process would be difficult in practice. As results illustrate, if the slope of the functions determining fuel economy targets dependent on vehicle footprint is flattened, the incentive to increase vehicle size is reduced. Results also suggest that, unless consumer preferences for vehicle size are at the lower bound and preferences for acceleration performance are at the upper bound of the ranges considered, the slope of both passenger car and light truck functions should be flattened and the slope of the function for light trucks should be flattened to a greater extent to avoid a divergence between the sizes of light trucks and passenger cars.

This analysis shows that designing the footprint-based CAFE standards such that no incentive exists to change vehicle size is complicated by the fact that this incentive depends on a number of relationships that vary among individual vehicle models. The incentive to increase vehicle size depends on engineering trade-offs between vehicle size and other vehicle attributes, consumer preferences for all of these attributes, production costs, and competition between automotive firms. Results illustrate that the incentive to change vehicle size resulting from these factors varies substantially across individual vehicle models. Consequently, designing footprint-based fuel-economy standards in practice such that manufacturers have no incentive to adjust the size of their vehicles appears elusive at best and impossible at worst.

6. Conclusions and recommendations

This study presents an oligopolistic equilibrium model to study whether footprint-based fuel-economy standards create an incentive to increase vehicle size. Simulation results reject the hypothesis that footprint-based standards do not create an incentive to increase vehicle size over a large range of assumptions regarding consumer preference. Except for the scenarios in which consumer preference for vehicle size is at the lower bound and preference for acceleration performance is at the upper bound of ranges considered, an incentive to increase vehicle size exists and can undermine gains in fuel economy. The required fuel-economy standards from these simulation results are 1.4–4.1 mpg (0.6–1.7 km/L) lower than if vehicle size and production mix is assumed unaffected by the policy. Results also suggest that the incentive to increase vehicle size is larger for light trucks than passenger cars, which could lead to higher traffic safety risks due to the increased divergence of vehicle size between these two classes. Furthermore, this analysis illustrates that incentives to change vehicle size vary considerably between individual vehicles, suggesting that modifying the CAFE standards as they are currently structured so that manufacturers do not have incentives to change the sizes of their vehicles is extremely difficult.

In the near-term, the analysis suggests that the following three measures could help to reduce the incentive to increase vehicle size. First, the slope of the function determining fuel economy targets based on vehicle footprint should be flattened for both passenger cars and light trucks, and even further for light trucks to avoid a divergence in size between these vehicle classes. Second, potential incentives for automakers to change vehicle size in response to the CAFE standards should be carefully analyzed in all future rulemakings to inform the specific policy design. Finally, considering the sensitivity of the incentive to increase vehicle size on consumer preferences, which are likely to change over time, future rulemaking should either allow for modifications to the standards if it becomes clear that fuel-economy goals will not be met or endeavor to design the standards such that the effects of changes in consumer preferences are minimized.

In the longer term, alternative policy options should be considered to address fuel-economy goals and concerns regarding traffic safety. The ideal solution would be a policy that could assess the impact of a vehicle on total traffic safety (including the vehicle's passengers, passengers of other vehicles, and pedestrians) as well as assess the impact of the vehicle on total fuel consumption and would optimize these two objectives for the social good, giving automakers guidance on how to balance the objectives where they compete and rewarding them for developing solutions that improve both safety and fuel economy. Considering the practical difficulties of designing and implementing safety and fuel-economy regulations, however, this ideal is clearly a long way off if not impossible. All the same, policymakers and researchers should consider how to make steps toward this ideal.

Acknowledgments

This research was supported by the Michigan Memorial Phoenix Energy Institute and the National Science Foundation MUSES program (CMMI 0628162). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the University of Michigan, or the National Academies. The authors would also like to thank Meredith Fowlie and John DeCicco for their helpful comments.

Appendix A

This appendix describes the regression analysis of the relationship between vehicle footprint and curbweight.

[3] This calculation is based on the EPA's estimates of 19.4 lb CO_2 per gallon gasoline and 22.2 lb CO_2 per gallon diesel (EPA, 2005), with gasoline vehicles making up 99% of new vehicle sales and diesel vehicles making up the remaining 1%.

Table A1
Estimates of the curbweight model in Eq. (A1).

Variable	I	II
logftp	1.3099243***	0.52676129***
height		0.00792866***
engsize		0.0613073***
_cons	3.1699523***	5.4764391***
N	472	472
r^2	0.68780556	0.80888083

*$p < 0.01$; ** $p < 0.001$; *** $p < 0.0001$.

The model of vehicle curbweight as a function of vehicle footprint is assumed to take the following form:

$$\log(wt) = \beta \log(ftp) + \gamma X + \varepsilon \tag{A1}$$

where wt is the curbweight of the vehicle, ftp is the footprint, $\mathbf{X}$ is a vector of covariates, and ε is the error term. The coefficient β is the percentage increase in curbweight resulting from a 1% increase in footprint (see for example Wooldridge 2002).

Two specifications of this model are used. The first uses no covariates; footprint is the only explanatory variable for curbweight. The second specification includes additional vehicle attributes as covariates to control for correlations in the data between footprint and other vehicle attributes that affect curbweight. Engine size (engsize) and vehicle height are included as covariates in this second specification.

Vehicle data from model-year 2006 was used to perform these regressions (Chrome Systems, Inc., 2008). Results of the three specifications are presented in Table A1. These results indicate that, when no additional vehicle characteristics are used as covariates, curbweight is estimated to increase by 1.3% for every 1% increase in footprint. When both engine size and height are controlled for in the regression, curbweight is estimated to increase by 0.53% for every 1% increase in footprint.

References

Austin, D., Dinan, T., 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.

Anderson, M., Auffhammer, M., 2011. Vehicle Weight, Highway Safety, and Energy Policy. University of California, Berkeley. Working Paper.

Anderson, S., Sallee, J., 2009. Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-economy Standards. Michigan State University. Working paper.

Beresteanu, A., Li, S., 2008. Gasoline Prices, Government Support, and the Demand for Hybrid Vehicles in the U.S. Duke University. Working paper.

Berry, S., 1994. Estimating discrete-choice models of product differentiation. RAND Journal of Economics 25 (2), 242–262.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica 63, 841–890.

Braess, H.-H., Seiffert, U. (Eds.), 2005. Handbook of Automotive Engineering, SAE International, Warrendale, PA.

Boyd, J.H., Mellman, R.E., 1980. The effect of fuel economy standards on the US automotive market: an hedonic demand analysis. Transportation Research Record A 14, 367–378.

Chrome Systems, Inc., 2008. Chrome New Vehicle Database, ⟨www.chrome.com⟩.

Clark, K., Fujimoto, T., 1991. Product Development Performance: Strategy, Organization, and Management in the World Auto Industry. President and Fellows of Harvard College, Cambridge, MA.

Fay, J.A., Golomb, D., 2002. Energy and the Environment. Oxford University Press, Oxford, UK.

Greene, D.L., Hopson, J.L., 2003. An analysis of alternative forms of automotive fuel economy standards for the United States. Transportation Research Board 1842, 20–28.

Greene, D.L., Keller, M., 2002. Appendix A: Dissent on safety issues : Fuel economy and highway safety. In: Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards. National Academy Press, Washington, DC.

Greene, D.L., Liu, J.T., 1987. Automotive fuel economy improvement and consumer surplus. Transportation Research A 22A (3), 203–218.

Greene, D.L., Patterson, P.D., Singh, M., Li, J., 2005. Feebates, rebates and gas-guzzler taxes: a study of incentives for increased fuel economy. Energy Policy. 33 (6), 757–775.

Goldberg, P.K., 1998. The effects of the corporate average fuel efficiency standards in the US. The Journal of Industrial Economics 46 (1), 1–33.

Helfand, G., Wolverton, A., 2009. Evaluating the Consumer Response to Fuel Economy: A Review of the Literature. National Center for Environmental Economics. Working Paper #09–04.

Jacobsen, M., 2010. Evaluating U.S. Fuel Economy Standards In a Model with Producer Household Heterogeneity. Stanford University. Working Paper.

Kahane, C.J., 1997. Relationships between vehicle size and fatality risk in model year 1985 – 93 passenger cars and light trucks. DOT HS 808 570. National Highway Traffic Safety Administration, Washington, DC.

Kelkar, A., Roth, R., Clark, J., 2001. Automobile bodies: can aluminum be an economical alternative to steel? Journal of Automotive Materials 53, 28–32.

Klier, T. Linn, J., 2008. New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard. Federal Reserve Bank of Chicago. Working Paper No. 2008–13.

McGee, J.S., 1973. Economies of size in auto body manufacture. Journal of Law and Economics 16 (2), 239–273.

National Research Council, 2002. Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards. National Academy Press, Washington, DC.

Nevo, A., 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. Journal of Economics and Management Strategy 9 (4), 513–548.

Norman, M.E., 1994. Reducing gasoline use: a multipronged approach. Energy Policy 22 (1), 37–39.

Sörenson, D., 2006. The Automotive Development Process: A Real Options Analysis. Deutscher Universitäts-Verlag, Frankfurt, Germany.

Stodolski, F., Vyas, A., Cuenca, R. Gaines, L., 1995. Life-cycle energy savings potential from aluminum intensive vehicles. Argonne National Laboratory.

Society of Automotive Engineers (SAE) International, 2005. Automotive Engineering Handbook. SAE International, Warrendale, Pennsylvania.

Train, K., Winston, C., 2007. Vehicle choice behavior and the declining market share of U.S. automakers. International Economics Review 48 (4), 1469–1496.

U.S. National Highway and Traffic Safety Administration, 2005. Proposed Rules, Average Fuel Economy Standards for Light Trucks, Model Years 2008 – 2011. Federal Register, vol. 70, no. 167, August 30.

U.S. National Highway and Traffic Safety Administration, 2006. Average Fuel Economy Standards for Light Trucks, Model Years 2008 – 2011. Federal Register 49 CFR 523-537, May 22.

U.S. National Highway and Traffic Safety Administration, 2008. Preliminary Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2011 – 2015 Passenger Cars and Light Trucks. U.S. Department of Transportation.

U.S. National Highway and Traffic Safety Administration, 2009. Proposed Rulemaking to Establish Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards. Federal Register 74, no. 186 (September).

Weber, J., 2009. Automotive Development Processes: Processes for Successful Customer Oriented Vehicle Development. Springer, New York.

Whitefoot, K. S., Fowlie, M. Skerlos, S. J., 2011. Product Design Responses to Industrial Policy: Evaluating Fuel Economy Standards using an Engineering Model of Endogenous Product Design. Energy Institute at Haas Working Paper WP-214. http://ei.haas.berkeley.edu/pdf/working_papers/WP214.pdf.

For the specialism to emerge and grow, data scientists will have to overcome barriers that are common to multidisciplinary research. As well as acquiring understanding of a range of science subjects, they must gain academic recognition. Journals such as the *Data Science Journal* should become more prominent within the computing community. Software products and technologies should be valued more by academic committees.

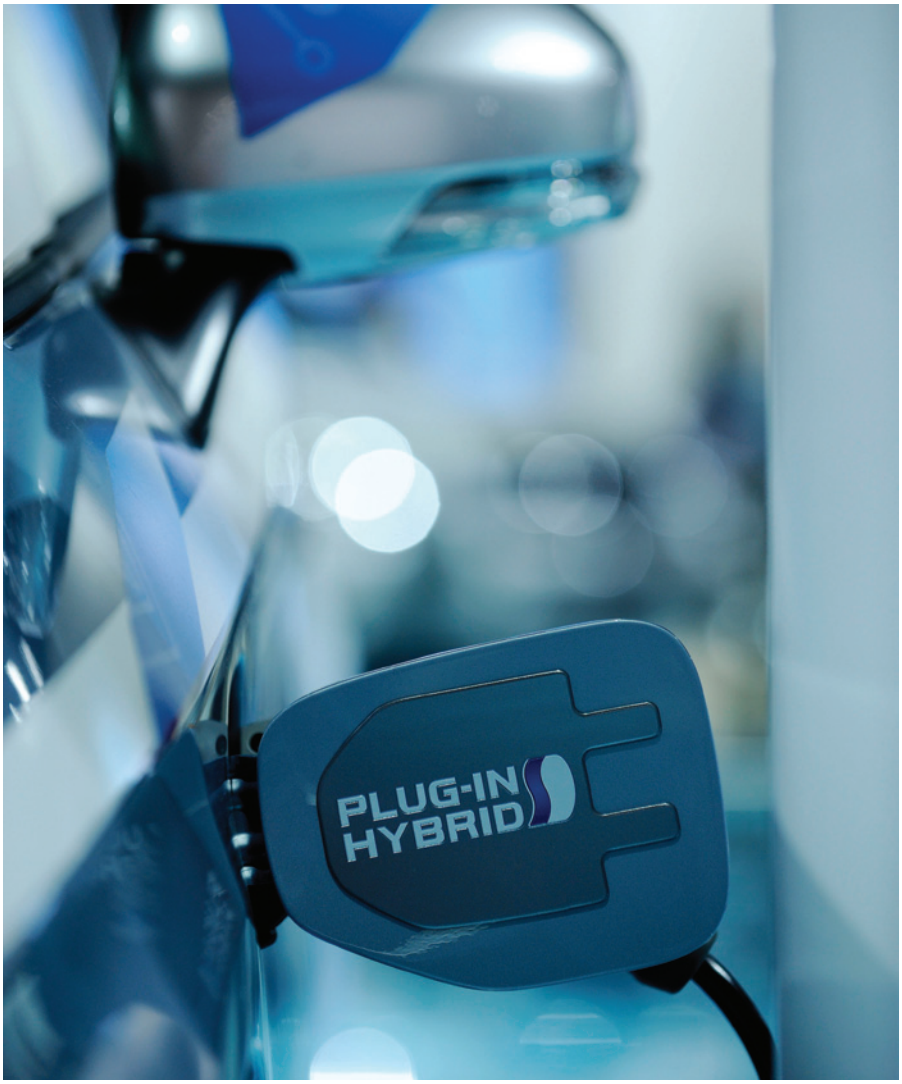
New interdisciplinary courses will be needed. The University of California, Berkeley, and Stanford University in California have set up introductory courses for computer scientists on big-data techniques — more universities should follow suit. Natural scientists, too, should become familiar with computing and format issues.

In my lectures for computer-science graduates, I have brought together students at the University of Southern California in Los Angeles with researchers at the JPL. Using real projects, my students see the challenges awaiting them in their future careers. I hope to employ some of them on the projects that will flow from the JPL's big-data initiative. The technologies and approaches that they develop will spread beyond NASA through contributions to the open-source community.

Empowering students with knowledge of big-data infrastructures and open-source systems now will allow them to make steps towards addressing the major challenges that big data pose. ■

Chris A. Mattmann *is a senior computer scientist at the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA, and adjunct assistant professor in computer science at the University of Southern California, Los Angeles, California 90089, USA.*
e-mail: chris.a.mattmann@nasa.gov

1. Painter, T. H., Bryant, A. C. & Skiles, S. M. *Geophys. Res. Lett.* **39,** L17502 (2012).
2. Foster, I., Kesselman, C. & Tuecke, S. *Int. J. High Perform. Comput. Appl.* **15,** 200–222 (2001).
3. Lynch, C. *Nature* **455,** 28–29 (2008).
4. Morin, A. *et al. Science* **336,** 159–160 (2012).
5. Spinellis, D. & Giannikas, V. *J. Syst. Softw.* **85,** 666–682 (2012).
6. Ven, K., Verelst, J. & Mannaert, H. *IEEE Software* **25,** 54–59 (2008).
7. White, T. *Hadoop: The Definitive Guide* 2nd edn (O'Reilly Media/Yahoo Press, 2010).
8. Mattmann, C. A. & Zitting, J. L. *Tika in Action* (Manning, 2011).
9. Cinquini, L. *et al. Proc. 2012 IEEE 8th Int. Conf. E-Science* Chicago, Illinois, 8–12 October 2012 (in the press).
10. Mattmann, C. A., Crichton, D. J., Medvidovic, N. & Hughes, S. in *Proc. 28th Int. Conf. Software Engineering (ICSE06), Software Engineering Achievements Track* 721–730 (2006).



ZHANG JUN/XINHUA PRESS/CORBIS

Fuel-efficient cars cost less to run, so people might use them a little more.

The rebound effect is overplayed

Increasing energy efficiency brings emissions savings. Claims that it backfires are a distraction, say **Kenneth Gillingham** and colleagues.

Buy a more fuel-efficient car and you will spend more time behind the wheel. That argument, termed the rebound effect, has earned critics of energy-efficiency programmes a voice in the climate-policy debate, for example with an article in *The New York Times* entitled 'When energy efficiency sullies the environment'[1].

The rebound effect idea — and its extreme variant the 'backfire' effect, in which supposed energy savings turn into greater energy use — stems from nineteenth-century economist Stanley Jevons. In his 1865 book *The Coal Question*, Jevons hypothesized that energy use rises as industry becomes more efficient because people produce and consume more goods as a result[2].

The rebound effect is real and should be considered in strategic energy planning. But it has become a distraction. A vast ▶

academic literature shows that rebounds are too small to derail energy-efficiency policies. Studies and simulations indicate that behavioural responses shave 5–30% off intended energy savings (see 'Bounce back'), reaching no more than 60% when combined with macroeconomic effects.

There is ample scientific evidence to diminish undue concern about rebounds and bolster support for energy-efficiency measures.

Many countries are considering legislation to limit energy demand, oil imports and pollution[3]. China plans to reduce its energy intensity by 16% from 2010 levels by 2015; the European Union aims to cut energy use by 20% compared with 2020 projections; and Japan seeks a 10% drop in electricity demand from 2010 levels by 2030. Energy efficiency could contribute to the savings, but no country is taking full advantage of its potential.

Various factors slow the uptake of efficient technologies, including behaviour, high cost and split incentives between investors and beneficiaries. Energy standards could help. Last year, the United States extended its fuel-economy standards for cars and trucks to require a doubling by 2025. Even taking rebound into account, we expect that these standards will yield substantial net energy savings.

FOUR EFFECTS

A rebound effect manifests in four ways, each of which makes energy-efficiency policies less effective. The important question is by how much.

The 'direct' effect occurs when a drop in the price of using an energy service causes a rise in demand. Analysts infer the size of the effect from changes in people's behaviour as prices vary. Numerous studies show that increased driving due to improved fuel economy reduces intended energy savings by 5–23% at first, rising to around 30% after several years as people get used to the lower cost[4]. The initial direct effect for home electrical appliances is also around 10% (ref. 5).

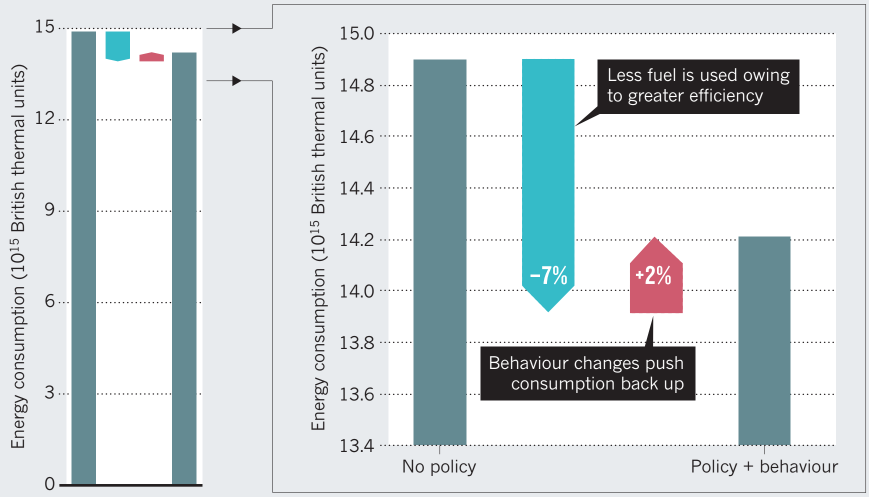
> ***"Rebound effects are small and are therefore no excuse for inaction."***

Because people respond more strongly to price than to efficiency cues when deciding how much energy to use[6], these numbers are overestimates. The direct rebound effect for efficiency alone should be nearer the low end of this range, or around 5–10% (refs 4,5).

Money saved through efficiency can also be spent on another product, such as a new phone, causing an 'indirect' rebound effect if extra energy is needed to manufacture and use the additional item. Assessments of household spending indicate that 5–15% of energy-efficiency savings are displaced in this way[7]. If the cost of making efficiency improvements is included, then the indirect effect is at the low end of this range. A Toyota Prius, for example, is more expensive than a comparable but less-efficient car, reducing the spare money available.

BOUNCE BACK

US policy to double fuel economy standards by 2025 would reduce projected energy consumption by almost 7% in the absence of behaviour changes. A conservative estimate of rebound effects from driving more and purchasing additional goods limits energy reductions to around 5%.



DATA FROM *EIA ANNUAL ENERGY OUTLOOK 2012*/AUTHORS' CALCULATIONS

Two other rebound effects apply on the scale of national economies. The latest fuel-economy standards passed by the United States will reduce demand for oil there. But, because that will drive down the price of oil globally, they could encourage people elsewhere to drive more, leading to a 'macroeconomic price' effect.

Greater energy efficiency could also spur pockets of industrial growth, leading to a 'macroeconomic growth' effect. Higher energy efficiency in one sector can create opportunities or technologies in others that consume more energy. For example, the development of lighter, stronger materials for fuel-efficient cars might lead to better aeroplanes, boosting energy use in the aviation sector.

Macroeconomic rebound effects are hard to pin down, but simple economic theory sets a limit. Standard assumptions linking supply and demand suggest that 'backfire' due to the price effect is impossible: if global demand for oil falls, the oil will become cheaper, so the incentive to produce it will be reduced. Less oil will be used overall, even though the cost is lower.

COMPLICATED SUMS

The four rebound effects cannot simply be added together to give the combined effect, because the presence of one may erode others. For example, when both the direct and indirect apply, the result is less than the sum of the two because any direct rebound effect decreases the amount of money available to spend elsewhere. Macroeconomic models estimate total combined rebound effects to be in the range of 20–60%[8,9].

In sum, rebound effects are small and are therefore no excuse for inaction. People may drive fuel-efficient cars more and they may buy other goods, but on balance more-efficient cars will save energy.

Energy-efficiency measures should be on the policy menu to curb energy use and to address global warming. Stricter energy-efficiency legislation should be considered across all sectors, alongside options that are not subject to rebound effects, such as carbon pricing. ■

Kenneth Gillingham *and* **Matthew J. Kotchen** *are at the Yale University School of Forestry and Environmental Studies, Connecticut, USA.* **David S. Rapson** *is with the Department of Economics at the University of California, Davis, USA.* **Gernot Wagner** *is at the Environmental Defense Fund in New York, USA.*
e-mail: kenneth.gillingham@yale.edu

1. Tierney, J. When Energy Efficiency Sullies the Environment. *The New York Times* (7 March 2011).
2. Jevons, W. S. *The Coal Question* (Macmillan, 1865).
3. International Energy Agency *World Energy Outlook 2012* (OECD/IEA, 2012).
4. Gillingham, K. T. PhD thesis, Stanford Univ. (2011); available at go.nature.com/krwzxm.
5. Jessoe, K. & Rapson, D. *Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use* NBER Working Paper 18344 (2012).
6. Chetty, R., Looney, A. & Kroft, K. *Am. Econ. Rev.* **99,** 1145–1177 (2009).
7. Thomas, B. & Azevedo, I. *Ecol. Econ.* (in the press).
8. Barker, T., Ekins, P. & Foxon, T. *Energy Policy* **35,** 4935–4946 (2007).
9. Barker, T., Dagoumas, A. & Rubin, J. *Energy Efficiency* **2,** 411–427 (2009).

Sensitivity of Vehicle Market Share Predictions to Discrete Choice Model Specification

C. Grace Haaf
Mechanical Engineering,
Carnegie Mellon University,
Pittsburgh, PA 15213
e-mail: grace.haaf@gmail.com

Jeremy J. Michalek
Professor
Mechanical Engineering,
Engineering and Public Policy,
Carnegie Mellon University,
Pittsburgh, PA 15213
e-mail: jmichalek@cmu.edu

W. Ross Morrow
Assistant Professor
Mechanical Engineering,
Iowa State University,
Ames, IA 50011
e-mail: wrmorrow@iastate.edu

Yimin Liu
Technical Expert
Ford Motor Company,
Dearborn, MI 48121
e-mail: Yliu59@ford.com

When design decisions are informed by consumer choice models, uncertainty in choice model predictions creates uncertainty for the designer. We investigate the variation and accuracy of market share predictions by characterizing fit and forecast accuracy of discrete choice models for the US light duty new vehicle market. Specifically, we estimate multinomial logit models for 9000 utility functions representative of a large literature in vehicle choice modeling using sales data for years 2004–2006. Each model predicts shares for the 2007 and 2010 markets, and we compare several quantitative measures of model fit and predictive accuracy. We find that (1) our accuracy measures are concordant: model specifications that perform well on one measure tend to also perform well on other measures for both fit and prediction. (2) Even the best discrete choice models exhibit substantial prediction error, stemming largely from limited model fit due to unobserved attributes. A naïve "static" model, assuming share for each vehicle design in the forecast year = share in the last available year, outperforms all 9000 attribute-based models when predicting the full market one year forward, but attribute-based models can predict better for four year forward forecasts or new vehicle designs. (3) Share predictions are sensitive to the presence of utility covariates but less sensitive to covariate form (e.g., miles per gallons versus gallons per mile), and nested and mixed logit specifications do not produce significantly more accurate forecasts. This suggests ambiguity in identifying a unique model form best for design. Furthermore, the models with best predictions do not necessarily have expected coefficient signs, and biased coefficients could misguide design efforts even when overall prediction accuracy for existing markets is maximized.
[DOI: 10.1115/1.4028282]

1 Introduction

Design researchers have proposed a variety of methods to predict the influence of design decisions on firm profit as part of a broader effort to base design decisions explicitly on predictions of downstream consequences for the firm [1]. The majority of these methods apply discrete choice methods [2] to predict consumer choice as a function of product attributes and price. Such predictions are proposed as a way to guide or even optimize design decisions [3–11]. Application of choice models within design implicitly relies on accurate choice predictions [5,12]. Given the many sources of uncertainty in such models, however, Frischknecht et al. [8] question the suitability of using choice models in a design context. At a minimum, researchers must be aware of the degree of prediction error and uncertainty when employing market models in design.

Prediction error can arise from many sources, including noisy data, finite data, omitted variables, changes in preferences or market conditions between estimation and prediction, and misspecification of the choice process [13]. Recent design research has modeled some aspects of model uncertainty by posing distributions over model coefficients [5,12]. Following standard asymptotic results, coefficient distributions are most often assumed to be normal with mean vector and covariance matrix determined by properties of the log-likelihood function. However, model misspecification is virtually guaranteed in most revealed preference contexts, given the complexity of human choice behavior for difficult decisions [14], and standard statistical results do not apply in such settings, nor are they comprehensive. Moreover, few applications of choice modeling in any field carefully analyze sensitivity of model fit or forecast accuracy using alternative utility specifications or error structures that might imply different design decisions. A realistic portrait of these aspects of predictive error cannot be captured in a fully generalizable way across product domains or contexts but can nevertheless be better understood via data-driven examination in the specific market of interest.

We focus on the effect of model specification and characterize share prediction accuracy of multinomial logit models in an empirical study of recent new vehicle markets using revealed preference sales data. The automotive sector is among the most popular product domains for application of choice modeling in general [4,7–9,11,15–44] and in the design literature specifically [4,7–9,15,21,24,27,28,32,35]. Logit models, along with variants including nested and mixed logit models, represent the most popular modeling approach by far. While stated choice methods fit to conjoint survey data are common [3,9,24,27,39–41], they measure hypothetical choices and generally must be calibrated to achieve a match with market sales data [25,45]. We focus here on choice models fit to aggregate market sales data [4,7,8,15–20,22,28,29,32–38,40,43,46].

Given the importance of the vehicle choice application in the design literature and beyond, a better understanding and characterization of prediction accuracy in this domain and its implications for design is needed. We aim to address this need with an automotive case study by fitting a set of models representative of those in the literature to past vehicle sales data, using the resulting models to predict sales in later years, and assessing prediction accuracy.

Our analysis is focused on the following research questions:

(Q1) How should we measure prediction accuracy, and do different measures lead to different conclusions about which models predict best?
(Q2) How widely do predictions vary for alternative model specifications? Which specifications have the best predictions, and how good are they?
(Q3) What are the implications for using choice models in design, particularly of new products?

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received May 28, 2013; final manuscript received June 23, 2014; published online October 20, 2014. Assoc. Editor: Bernard Yannou.

The design literature has not yet investigated what measures of forecast accuracy exist or compared these measures to understand how they differ in characterizing accuracy, thus Q1. Q2 applies appropriate measures to the specific task in our case study. Q3 focuses on prediction accuracy for new vehicle designs, and we examine the relationship between accurate prediction in existing markets versus potential to predict response to new designs that deviate from market patterns (e.g., correlations with unobserved attributes). We view design as primarily interested in the introduction of new products or (large) changes to product features, motivating a focus on new vehicles.

2 Literature Review

Broadly, there are two schools of research in the vehicle demand literature. The first is concerned foremost with predicting future vehicle demand shares, usually at an aggregate level like vehicle class or powertrain type, and often without transparency about the assumptions and models used to make the forecast. We henceforth refer to this type of literature as "forecasting". The second school is interested in model construction and in vehicle and consumer attributes coefficient estimation especially as it pertains to willingness-to-pay and demand elasticity in past markets. We henceforth refer to this type of literature as "explanatory." Appendix A compares publications of each type.[1]

Forecasting studies are conducted by private or government research entities or issued in report format from an academic research institute (see Appendix A). Reports are typically not peer reviewed and rarely contain a full mathematical description of the model, making it impossible to reproduce the model without additional information. Some reports include sensitivity cases formed with variations on model assumptions; for example, the Energy Information Administration Annual Energy Outlook [47] contains base, low, and high alternative vehicle future market share as a result of base, low, and high future oil prices. This type of sensitivity only captures uncertainty about model input parameters and assumes that model specification and estimated coefficients are known. In practice, model specifications for choice contexts as complex as automotive purchases are always uncertain, and the relevant question is whether or not the model is sufficient for its intended function. The forecasting literature is typically not used in engineering design models due to lack of transparency and documentation of data and modeling assumptions and lack of models that make predictions as a function of design variables. Rather, models from the explanatory literature are applied in a predictive context.

The bulk of the new vehicle purchase demand literature is explanatory, conducted by academic researchers and published in peer-reviewed academic journals (see Appendix A). This literature extensively discusses model estimation and to a lesser degree model selection, including potential sources of error from model misspecification. Usually researchers compare the goodness-of-fit across several specifications in order to determine which model best represents a known, current reality. However, most of this literature does not attempt to make predictions about future vehicle market share penetration or evaluate models with predictive capabilities in mind (Frischknecht et al. [8] is a rare exception). In general, models that fit the existing data best may not necessarily be the best at predicting counterfactuals: statistical models may be misspecified, containing systematic difference in prediction from true process ("bias"), or may be sensitive to overfitting noise in the data instead of signal ("variance") [48].

The earliest applications of economic models for overall automotive demand focused on macroeconomic variables and, as Train [49] highlights, only included price. These studies are referred to as aggregate studies because the level of granularity of predictions is at the whole market or vehicle class level as opposed to individual vehicle designs[2]. Disaggregate studies evolved to predict the number of vehicles an individual household would choose to own [49]. For example, Lave and Train [44] advanced this work by proposing a disaggregate model of vehicle class purchase choice based on consumer characteristics and additional vehicle characteristics, such as fuel economy, weight, size, number of seats, and horsepower. A wide variety of models followed over the next three decades: Boyd and Mellman [43], who propose a random coefficient logit model adopted by others [11,28,35,50,51]; Berry et al. [42], who include an alternative-specific constant (ACS) in the utility function of a random coefficient demand model adopted by others [16,52–54]; Brownstone and Train [39], who propose several choice model specifications using the results of a California conjoint study described in Bunch et al. [55] and adopted by others [56,57]; and Whitefoot and Skerlos [11], who investigate the effect of fuel economy standards on vehicle size and employ a logit model with coefficients drawn directly from the literature. Other new-vehicle purchase models include [23,26,29,31,34,40,41,58].

We use the preceding literature to inform comparison models of our creation; we do not recreate prior models exactly due to limited availability of data or specifics about estimation methods. Instead, we form a combinatorial set of utility specifications using covariate forms from these prior models, fit them all to a common data set, and test them all on a common prediction set. Appendix B summarizes the covariates used in past models and those adopted for our tests.

3 Methods

Our overall goals are to examine the robustness of multinomial logit model predictions over various utility function specifications and to compare the predictions across the structural specifications of logit, mixed logit, and nested logit (for brevity we refer to the multinomial logit model as "logit"). We identify a universe of covariates informed by the literature and form combinations of them such that we have defined all possible linear utility function specifications from these covariates. We then estimate the logit coefficients on US consumer vehicle purchase data from 2004 to 2006 and predict market share for each of the vehicles in the US purchase data from 2007 and 2010.

Using the measures described in Sec. 3.4, we rank the predictive accuracy across utility function specification for each of the measures.

3.1 The Data Set. Our data set draws vehicle attribute information from Ward's Automotive Index [59] and aggregate US sales data from Polk [60] for vehicle sales during 2004–2007 and 2010. Other studies have used a variety of data sources (including these) as well as stated preference surveys. We use 2004–2006 data for estimation because we expect three years of data to be sufficient to predict a successive year, and we predict 2007 and 2010 sales to examine the effects of different time horizons. We implicitly assume that all individuals who purchased a vehicle considered all of the other vehicles available in the same year and made a compensatory decision based on vehicle attributes.

Our models consider only new vehicle buyers, thus there is no outside good (option to not purchase any vehicle). Inclusion of an outside good allows a choice model to endogenously determine market size. Excluding it models only share among the vehicles purchased, which is likely less sensitive to macroeconomic factors. There are many factors that drive share and are not included in our models, but we are interested in how well a modeler can predict when relying primarily on available vehicle attribute data.

3.2 Model Specification. Each model uses the utility function

[1]An electronic companion to this paper containing the appendices referenced herein can be found at http://repository.cmu.edu/meche/70/

[2]We use the term "vehicle design" to refer to vehicle make-model.

$$u_{ij} = \mathbf{x}_j' \boldsymbol{\beta} + \varepsilon_{ij} \tag{1}$$

where u_{ij} is the utility of vehicle design j for consumer i, $\mathbf{x}_j$ is the attribute vector of vehicle j, $\boldsymbol{\beta}$ is the vector of model parameters to be estimated, and ε_{ij} is an error term. Following standard assumptions, if ε_{ij} is independently identically distributed (iid) and follows a type I extreme value distribution, then the probability P_j that a randomly selected consumer will choose vehicle j can be expressed as

$$P_j = \frac{\exp\left(\mathbf{x}_j' \boldsymbol{\beta}\right)}{\sum_{k=1}^{J} \exp\left(\mathbf{x}_k' \boldsymbol{\beta}\right)} \tag{2}$$

where J is the number of vehicle design options. This is the (multinomial) logit formula.

While any choice of covariates $\mathbf{x}$ is possible in principle, we focus on combinations of covariates used in the prior literature. We survey the automotive demand literature to identify the universe of independent variables historically used in automotive discrete choice models (Appendix B). From this list of candidate covariates, we select a subset to define a manageable set of models. Many of the models in the literature include demographic or consumer usage covariates, but because Polk sales data [60] does not include individual-level choices, we ignore demographics. For some demographic information like gender or income an aggregate distribution over the US population is available, but because we do not know which consumers selected which vehicles, sampled consumer attributes are unlikely to accurately determine specific individuals' sensitivity to vehicle attributes. We omit several variables because they are not available in our data sources:

- *Indirect vehicle attributes* like consumer reports ratings for handling and safety—These would be unknown at the time of prediction.
- *Vehicle and battery maintenance costs*—These covariates are used primarily when predicting alternative vehicle share, and they will not vary substantially across conventional and hybrid powertrains.
- *Acceleration time* (*seconds*)—We indirectly test inclusion of acceleration through functions of horsepower and weight. Note that horsepower/weight correlates well with 0–60 mph acceleration time for cars well but poorly for trucks.
- *Range*—This covariate is used primarily when predicting alternative vehicle share and will not vary substantially across conventional and hybrid powertrains. A related fuel economy covariate is included.
- *Top speed*—We use an alternative measure of performance through horsepower and weight.
- *Number of seats*—We use vehicle class, which is closely related to seating.
- *2-year retained value*—Like the consumer rating data this would not be known at the time of prediction.
- *Attributes specific to alternative-vehicles* (e.g., dummies for hybrid or electric power trains)—These are not relevant to our data set, which includes conventional vehicles and only a limited number of hybrid powertrains.

The highlighted covariates in Appendix B are those which remain after omitting demographic, usage, indirect, and unavailable attributes. Some studies group price and fuel economy variables into discrete levels of each rather than treating them as continuous variables. We consider all covariates (except for class and brand dummies) to be continuous variables because, unlike controlled conjoint experiments, the market data do not fit well into a small number of discrete levels. Price is always included as a covariate and can take any of the forms listed in Table 1; vehicle class dummies are also always included. The other highlighted covariates in Appendix B can take one of the forms listed in Table 1 or can be excluded from the utility function entirely ("excluded" option). Given these covariate options, there are 9000 possible utility specifications for the logit model outlined in Table 1. Operating cost includes the macroeconomic variable of retail gas price. Though we aim to exclude nonvehicle attributes, this covariate was particularly prevalent in the literature. Furthermore, while having more covariates cannot decrease best model fit on a given data set, that does not imply that more covariates will improve model forecast accuracy. In general, introducing more covariates introduces the risk of overfitting the estimation data.

From the selected covariates, we assume that the utility function is linear in parameters (a standard assumption in the vast majority of logit model applications because it ensures that the log-likelihood function is concave [2]) and construct models using all possible linear combinations of covariates.

Many of these covariates are correlated. Such correlations can induce bias in the estimated coefficients if not corrected [63]. However, while this presents difficulties in drawing inferences from the coefficients (e.g., willingness-to-pay) it does not necessarily affect the ability to make predictions from the model so

Table 1 Covariate forms tested in utility function specifications

	Functional form options				
Covariate	Option 0	Option 1	Option 2	Option 3	Option 4
Price		Price ($)	Price + op cost	ln(price)	
Operating cost[a]	Excluded	Fuel cost/mile	Miles/fuel cost	Miles/gallon	Gallons/mile
Acceleration[b]	Excluded	Horsepower/weight (hp/wt)	wt/hp	$\exp(c_1 \times (\mathrm{hp/wt})\, c_2)$	hp
Size	Excluded	Length	Width	Length-width	Length × width
Style	Excluded	(Length × width)/height			
Air conditioning	Excluded	Dummy if air-conditioning is standard			
Transmission	Excluded	Dummy if auto. transmission is standard			
Brand	Excluded	Dummy for country of origin[c]	Dummy for brand[d]		
Vehicle class		Dummies for vehicle class[e]			

[a]Fuel cost is average annual gas price [61] in 2004 dollars, adjustment based on the consumer price index [62].
[b]$c_1 = -0.00275$ and $c_2 = -0.776$ as in the EIA Annual Energy Outlook [47].
[c]Country of origin includes: United States, Europe, and Asia; excludes United States dummy for identification.
[d]Brand includes: Acura, Audi, BMW, Buick, Cadillac, Chevrolet, Chrysler, Dodge, Ford, GMC, Honda, Hummer, Hyundai, Infiniti, Isuzu, Jaguar, Jeep, Kia, Land Rover, Lexus, Lincoln, Mazda, Mercedes, Mercury, Mitsubishi, Nissan, Oldsmobile, Pontiac, Porsche, Saab, Saturn, Scion, Subaru, Suzuki, Toyota, Volkswagen, Volvo; excludes Acura dummy for identification.
[e]Class includes: Compact, midsize sedan, full size sedan, luxury sedan, SUV, luxury SUV, pickup, minivan, van, and sports; van is excluded for identification.

long as the correlations in the training data would also be present in the prediction set. For vehicle markets, this is likely to hold for near-term predictions, though it may not hold for new designs that do not follow prior patterns in the marketplace.

For illustration of this concept, suppose the true choice generator uses the utility function $u(\mathbf{x}|\boldsymbol{\beta}_0) = \boldsymbol{\beta}_0'\mathbf{x} + \varepsilon$, and the designs in the market follow a pattern: $\mathbf{x} = \mathbf{A}'\mathbf{y}$ for $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, m < n$. Then for any coefficient vector $\{\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\Delta} : \mathbf{A}\Delta = 0\}$, $u(\mathbf{x}|\boldsymbol{\beta}) = \boldsymbol{\beta}'\mathbf{A}'\mathbf{y} + \varepsilon = (\mathbf{A}\boldsymbol{\beta}_0 + \mathbf{A}\boldsymbol{\Delta})'\mathbf{y} + \varepsilon = \boldsymbol{\beta}_0'\mathbf{x} + \varepsilon = u(\mathbf{x}|\boldsymbol{\beta}_0)$. Therefore, choice probabilities are identical for any $\boldsymbol{\Delta}$ in the null space of $\mathbf{A}$, and $\boldsymbol{\beta}_0$ is not identifiable: coefficient estimates $\boldsymbol{\beta}$ could be arbitrarily far from their true value $\boldsymbol{\beta}_0$. Nevertheless, $u(\mathbf{x}|\boldsymbol{\beta}) = u(\mathbf{x}|\boldsymbol{\beta}_0)$, so utility estimates (and therefore choice probabilities) can be correct even for arbitrarily biased coefficients as long as the new designs follow the pattern in the marketplace $\mathbf{x} = \mathbf{A}'\mathbf{y}$. If a new design deviates from the prior pattern $\tilde{\mathbf{x}} = \mathbf{A}'\mathbf{y} + \mathbf{z}$, utility (and therefore choice probabilities) may be biased: $u(\tilde{\mathbf{x}}|\boldsymbol{\beta}) = (\boldsymbol{\beta}_0 + \boldsymbol{\Delta})'(\mathbf{A}'\mathbf{y} + \mathbf{z}) + \varepsilon = (\mathbf{A}\boldsymbol{\beta}_0 + \mathbf{A}\boldsymbol{\Delta})'\mathbf{y} + (\boldsymbol{\beta}_0 + \boldsymbol{\Delta})'\mathbf{z} + \varepsilon = \boldsymbol{\beta}_0'\mathbf{A}'\mathbf{y} + \boldsymbol{\beta}_0'\mathbf{z} + \boldsymbol{\Delta}'z + \varepsilon = u(\tilde{\mathbf{x}}|\boldsymbol{\beta}_0) + \boldsymbol{\Delta}'z$. Therefore, models that predict well overall may nevertheless have biased coefficients that predict poorly for new designs that deviate from the market pattern. We assess predictive accuracy for products in the marketplace and also examine variation in implications of coefficient estimates for new designs.

3.3 Model Estimation. The likelihood of the estimated parameters L is defined as the probability of generating the observed data given the estimated parameter values

$$L\left(\hat{\boldsymbol{\beta}}|\mathbf{x}\right) = \prod_{j=1}^{J} \left(P_j\right)^{n_j} \tag{3}$$

where n_j is the sales of vehicle j. The maximum likelihood estimator of the parameters $\hat{\boldsymbol{\beta}}$ is the value of the vector that maximizes L. The monotonic transformation $\ln(L)$ is typically used as the objective function for computational benefit. For more detail on logit models and their estimation see Train [2].

The mixed logit, or random coefficients logit, model is similar to the logit model except the individual β's are allowed to vary over the population to represent heterogeneous consumer preferences. In our case we assume that they are independently normally distributed

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{4}$$

where $\boldsymbol{\Sigma}$ is a diagonal matrix, and the maximum likelihood procedure estimates the elements of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using numerical integration [2]. This specification relaxes the independence from irrelevant alternatives (IIA) restriction for substitution patterns [2].

Our nested logit specification divides the vehicles into groups or nests by vehicle class and fits a logit model to each of the nests. We assume that the utility functional form is the same for each nest, but coefficients may differ across nests. For example, the β for price will be different for midsize cars than it is for pickups. However, within a nest β is fixed. A nested logit exhibits the IIA property for products within a nest, but relaxes the IIA restriction for products in different nests.

As generalizations of the logit model, nested and mixed logit models will necessarily fit any set of *estimation* data at least as well as the logit. The mixed logit generalization of the logit model is even flexible enough to represent most random utility maximization models, given enough flexibility over the coefficient distribution [2]. However, nested and mixed logit models need not *predict* as well as logit models due to the potential for overfitting.

3.4 Evaluation Measures. After fitting each of the model specifications, we evaluate prediction error using likelihood measures, the Kullback–Leibler divergence (KL) [64], a cumulative distribution of error tolerance (CDFET), and the average share error (ASE), and we compare the goodness-of-fit using the above measures as well as the Akaike information criterion (AIC) [65], and the Bayesian information criterion (BIC) [66]. Each of these measures is described below. We compare models selected as best by these measures to one another and to literature-informed benchmark models.

Likelihood: Likelihood, defined in Eq. (3), and monotonic transformations of likelihood, such as log-likelihood $\ln(L)$ and average likelihood (AL) ($L^{1/N}$, where N is the number of choices observed) measure the probability that the model would generate the data observed. When comparing two models for the same data set, the model with larger L is more likely to generate the data observed.

KL divergence: The KL divergence measures the difference between a predicted distribution and the true distribution [67].

$$KL\left(s_j||P_j\right) = \sum_{j=I}^{J} \ln\left(\frac{s_j}{P_j}\right) s_j \tag{5}$$

where $s_j = n_j/J$ is the market share of vehicle design j. The KL measure is also a monotonic transformation of L, thus L and KL will rank models identically, and maximizing likelihood is equivalent to minimizing KL (see Appendix C for proof).

ASE: ASE measures the average error in share predictions across the vehicle designs.

$$\mathrm{ASE} = \frac{1}{J}\sum_{j=1}^{J} \left|s_j - P_j\right| \tag{6}$$

We report ASE as a summary statistic in Appendix D but do not use it as a basis for model selection because it does not holistically capture distribution divergence: It will not distinguish between models with large error for one vehicle alternative vs. the same degree of error spread out among many vehicle alternatives.

Error tolerance cumulative distribution function (CDF): The CDFET graphs the fraction of vehicles with absolute share prediction error, $|s_j - P_j|$ for vehicle design j, less than a specified value. This measure, to our knowledge proposed here, evaluates a model in terms of error tolerance levels. We use absolute share error rather than relative error because relative error overemphasizes small prediction errors for vehicles with small market share. A CDFET is a more comprehensive description of model prediction error than likelihood measures because it characterizes the distribution of accuracy across the vehicle share predictions, rather than just how well a model predicts "on average".

Two additional measures apply only to assess fit with estimation data, not predictive accuracy [68].

AIC: AIC is a variation of likelihood that attempts to penalize overfitting.

$$\mathrm{AIC} = 2\ln(L) - 2k \tag{7}$$

where k is the number of model parameters.

BIC: BIC is similar to AIC but with a stronger penalty for an increasing number of covariates.

$$\mathrm{BIC} = 2\ln(L) - \ln(J)k \tag{8}$$

AIC and BIC can take on the value of any negative real number, have no standalone meaning, and are only useful as compared to other candidate models fit to the same data set. Larger values are preferred. Derivations and consistency proofs for the KL, AIC, and BIC measures can be found in Ref. [68].

4 Results

Of the 9000 tested utility function specifications, for 8993 (99.9%) the Knitro optimization algorithm for MATLAB converged

Table 2 RAL calculated on the prediction data set for select model specifications and data sets

Scenario	1	2	3	4	5	6
Estimation data	2004–2006	2006	2007	2004–2006	2004–2006	2004–2006
Prediction data	2007	2007	2007	2010	2007	2007
Market	Full market	Full market	Full market	Full market	Luxury sedan	New designs
AL of ideal model (predicted shares = actual shares)	*0.0076*	*0.0076*	*0.0076*	*0.0080*	*0.0384*	*0.4610*
RAL of no info model	55.3%	55.3%	55.3%	43.6%	63.6%	93.2%
RAL of static model	**88.3**%	**88.3**%	**88.3**%	23.7%	73.3%	95.9%
RAL of class dummies only logit	65.9%	65.9%	65.9%	53.6%	NA	95.0%
RAL of best fit logit model for L/AIC/BIC of estimation data	76.4%	77.7%	81.7%	67.3%	73.3%	96.5%
RAL of logit model with greatest likelihood for prediction data	79.0%	79.0%	81.7%	**68.5**%	87.9%	**97.5**%
RAL of mixed logit with best logit estimation fit covariates	79.0%	79.0%	85.6%	67.3%	**89.2**%	97.0%
RAL of nested logit with best logit estimation fit covariates	73.8%	72.4%	80.3%	64.8%	NA	95.3%

Note: in Scenario 5 luxury sedan vehicles are used for estimation and prediction; in scenario 6 the full market is used for estimation, but evaluation measures are assessed for prediction of new vehicles only. Italicized numbers emphasize that the number is an AL as opposed to an RAL. Bold numbers indicate the greatest RAL in a given column (most accurate model for a given scenario).

to likelihood-maximizing coefficients, and the other seven failed to converge. Only the 8993 models that successfully converged were considered as candidate models. The candidate models were ranked from best to worst on each measure. There were no two models with identical values for any measure (no ties). In the following results "best models" refer to the models ranked as number one for a given measure.

4.1 Q1: Model and Evaluation Measure Comparison. We refer to a model that most accurately predicts the in-sample estimation data according to a given measure as the "best estimative model", and we refer to a model that most accurately predicts the out-of-sample prediction data as the "best predictive model." The traditional goodness-of-fit measures—likelihood/KL and AIC/BIC—select the same best estimative model, and they also agree upon the specification of the best predictive model. The CDFET goodness-of-prediction measure selects distinct model specifications as the best predictive models dependent upon the desired error tolerance level (we test error tolerance levels of 25%, 50%, and 75%). The three CDFET best predictive models are also distinct from the best estimative and predictive models under the AIC, BIC, and likelihood criteria. See Appendix D for selected model measure comparisons and coefficient estimates.

Though the best likelihood/AIC/BIC estimative model is distinct from the best predictive model, the difference in form is small. They include the same covariates but in different forms (e.g., operating cost as miles/dollar as opposed to gallons/mile) with the exception of luxury and transmission which contribute little to utility relative to the contribution of the other attributes.

4.2 Q2: Model Accuracy. Table 2 summarizes the AL calculated on the prediction data set for select combinations of model specification (rows) and estimation/prediction data set scenarios (columns). We report the *relative* average likelihood (RAL) in Table 2 defined as the AL of the model divided by the AL of an ideal aggregate model that predicts shares perfectly. The reason we report RAL instead of simply AL is because choice diversity in the data necessarily lowers the maximum attainable value of AL with any model. Thus RAL describes the amount of predictive power obtained by a particular model relative to the best possible predictive power that could be obtained with any aggregate model.

The rows compare the predictive performance of the model that has the best predictions and the model that fit the estimation data best. Using each of the utility functions from the best estimative logit models, we fit additional mixed and nested logit models. Due to computational limitations, we did not run all 9000 utility form combinations for the mixed and nested logit structural specifications. Rather we used the results from the logit model output to inform the selection of covariate form for the mixed and nested logit models. The "no info" row is calculated by assigning an equal share to all vehicles. The static model row assumes that shares in the prediction year are identical to the most recent share of the vehicle design available in the estimation data for all vehicle designs present in both the estimation and prediction data, and all new vehicle designs receive an equal proportion of the remaining share.

Scenario 1 is our base case, where models are fit to sales in years 2004–2006 and used to predict 2007 sales. Scenario 2 uses only 2006 data to predict 2007, assessing sensitivity of predictions to the amount of data used for estimation. Scenario 3 fits the models directly to 2007 data, helping to identify the portion of prediction error that stems from model fit, rather than from changes over time. Scenario 4 uses 2004–2006 data to predict 2010 sales, assessing differences when predictions are made farther into the future. Scenario 5 assesses predicative accuracy for a single vehicle class,[3] rather than the entire market, and scenario 6 assesses only the predictive accuracy for new vehicle designs introduced in 2007. Comparisons can be made within each column to evaluate the prediction accuracy across model specifications for a given estimation/prediction data set.

In scenarios 1–3, which predict the full 2007 market, the best predictive logit model predicts better than the best estimative model, the class dummies model does not predict as well as the models which contain vehicle attributes, and the no info model predicts worst, as expected. Nested logit predictions have lower AL than logit, but mixed logit predictions have higher AL.[4] That the nested logit does not predict better than the logit suggests that the relaxation of the IIA property among the nests selected does not improve prediction. Model predictions could potentially be improved further by exploring alternative parameter distributional forms such as multivariate normal with a full covariance matrix [69], although that introduces more potential for overfitting with aggregate sales data. We leave such explorations for future work. See Appendix E for mixed and nested logit coefficient estimates and Appendix F for actual versus predicted shares.

In all three scenarios the static model outperforms all other models. Additionally, we see little difference in prediction quality between scenarios 1 and 3 when using the same model (compare across columns) compared to the difference due to model specification (compare down rows), even though the prediction set and

[3]This is distinct from the "class dummies only logit" which includes data for the entire market but uses only dummies representing each class as covariates.

[4]A likelihood ratio test of the best logit and mixed logit models calculated on 2007 data suggests that there is sufficient evidence to reject the null hypothesis that the mixed logit model predicts significantly better at the $\alpha = 0.1$ level.

estimation set are identical in scenario 3. Together, these results indicate that residual error in model fit is a major source of prediction error, and there is too much missing data or model misspecification in the attribute-based models to fit or predict the full market as well as the static model. Without data on missing covariates that influence choice, such as vehicle aesthetics, it is difficult to fully explain choice behavior at the vehicle design level with only the available covariates.

However, scenario 4 examines a longer time horizon and reveals that the static model has poor predictive capability when forecasting farther into the future. The attribute-based models attempt to capture consumer choice as a function of observable attributes plus random noise, but since not all attributes are observed, share is not fit perfectly. In contrast, the static model does not attempt to explain the reason for consumers' choices but instead simply assumes consumers will make the same choices year after year. The static model does well for the 2007 forecasts because share for each vehicle model changes little from year to year, but over a longer time horizon vehicle designs change and new designs are added to the market (~37% of the vehicle designs sold in 2010 did not appear in the 2004–2006 data). The static model has no information about these new designs, so it loses predictive capability, and over a longer prediction horizon the attribute-based models perform substantially better than the static model.

Scenario 5 indicates that the attribute-based models also perform better than the static model in the luxury sedan class. The best class model is distinct for each class, though all class models include some form of all covariates with the exception of style and automatic transmission as standard. The AL of 2007 class predictions increases when the best estimative class level model is fit to class data as opposed to the best estimative full market model fit to the full market data with the exception of midsize and sports cars (see Appendix G for table of class model specifications and model AL comparison by class).

Figure 1(*a*) shows the CDFET for selected models of scenario 1. The *x*-axis is the absolute difference between the predicted share and the actual share, and the *y*-axis is the proportion of vehicle designs whose share prediction error is less than the corresponding value on the *x*-axis. For example, in Fig. 1(*a*) point (0.25%, 0.7) indicates that 70% of the share predictions made by the best AIC/BIC/KL models deviate from the observed share by less than 0.25% (the average vehicle design share in this market is 0.42%).

The worst models all perform similarly to one another in scenario 1 and lie on top of the class-only curve in Fig. 1(*a*) (and are thus omitted for readability). While a model could plausibly be posed that predicts worse than the no info model, we do not observe it in our utility specifications. The best models and worst models differ most noticeably in their omission of covariates. The best models include some form of almost every covariate, whereas the worst models omit covariates entirely. For example, the worst model as selected by the likelihood and AIC measures applied to the estimated data only contains the covariates price and class. Conversely, if we compare only models that contain some form of price, operating cost, acceleration, size covariates, and class and brand dummies (style, luxury, and automatic transmission dummies could be excluded), then we see no practical difference in the predictive power of the best and worst models. No one covariate in isolation sets the best models apart from the worst models. A model's predictive power thus appears to be robust to covariate form but sensitive to the exclusion of attributes.

4.3 Q3: Implications for Design.

Scenario 6 compares the best-predictive logit model for all vehicles to the model that best predicts the shares of the new vehicle designs introduced in 2007. The best new vehicle model is determined similarly to the best predictive logit model of scenarios 1–3 by ranking the models on each of the measures; however, the measures in this case were calculated by treating each of the new vehicles individually and the holdover vehicles as an aggregated "other" share. (The "other" share is calculated as the sum of all holdover vehicle shares.) In contrast to scenario 1, the attribute-driven logit models of scenario 6 have a higher likelihood than the static model, since the static model has no information about new designs.

The CDFET of Fig. 1(*b*) shows that at lower values of error tolerance the attribute-driven models are superior to the static model and that there is some difference in prediction quality between models that predict best for the whole market versus the new vehicle market. Overall, while the static model outperforms attribute-based models for near-term predictions, attribute-based models are needed for predicting the performance of new vehicle designs and for making longer-term predictions. Still, the degree of uncertainty and error in predictions for new designs may be too large to guide design choices appropriately in some contexts.

Appendix D summarizes model coefficients for several specifications including those representative of models in the literature as well as best estimative and best predictive models. It is clear that different specifications lead to different inferences about the effect of attribute changes on choice. For instance, the utility function specifications based on Boyd and Mellman [43], Berry et al.

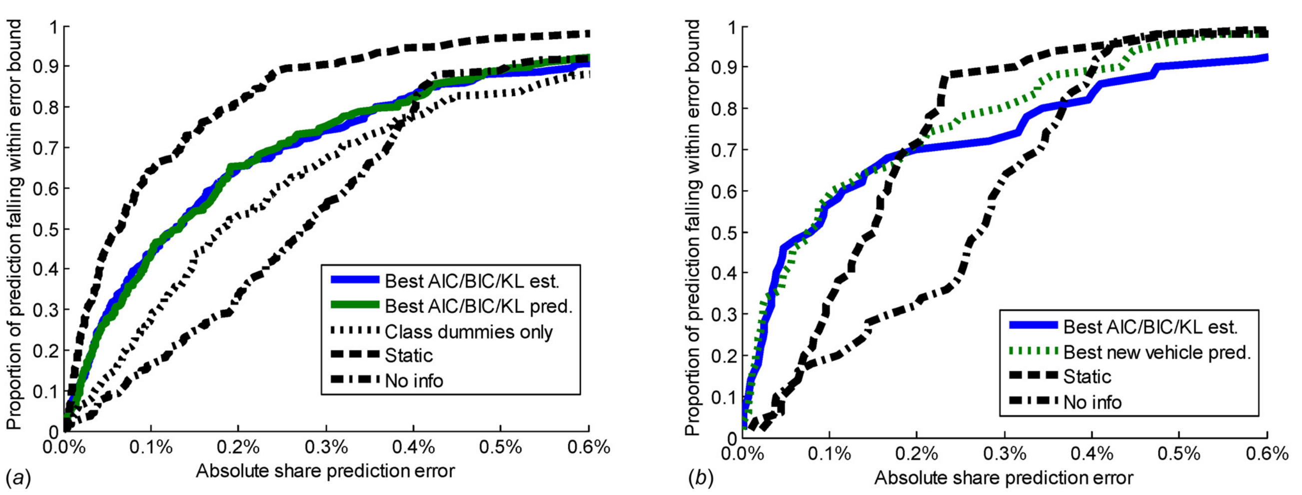


Fig. 1 CDF of error tolerance for the best logit model specifications as measured by likelihood/KL and AIC/BIC measures on 2004–2006 sales estimation data and 2007 sales prediction data compared to alternative models (*a*) full market and (*b*) new vehicle designs only

[42], and Whitefoot and Skerlos [11] result in a coefficient for operating cost that suggests consumers prefer higher efficiency (longer range per unit cost or lower fuel consumption per unit distance) all else being equal, as expected. But the best estimative and best predictive models suggest that consumers prefer lower efficiency. This can happen because efficiency may serve as a proxy for unobserved variables (e.g., size, performance, or styling variables not captured in the data). While the latter models make better predictions for existing vehicle markets that follow established patterns (attribute correlations), they could misguide design efforts that divert from established market patterns.

5 Limitations

Our investigation is a first step in a larger goal of characterizing the design impacts of choice prediction uncertainty. All of our models have error resulting from misspecification and missing information (as do all similar models in the literature that are based on market sales data rather than controlled experiments). For example, we do not have information on attributes that are important in some vehicle classes (like towing capacity for trucks), and we lack information and quantification of some key purchase drivers, such as esthetics. We lack individual-level choice data with consumer covariates, such as demographics or usage variables [9], which can help explain choice behavior and improve predictions when predictions of future population covariates are available. Nevertheless, such limitations are common in choice models used to assess the vehicle market or guide design choices. Our study suggests that if models lack transparent quantifications of important determinants of product choices, designers should be cautious about basing design decisions on choice models.

More research is needed to assess a wider scope of modeling alternatives. We did not consider ASCs—product-specific factors that can proxy for omitted variables—and their use in prediction or design. ASCs can generate models that match estimation data shares exactly; however, they contain no information about specific unobserved product features, and they are unknown for any new product designs. We also ignore a major component of the new vehicle modeling literature: covariate endogeneity—a correlation between model covariates and the unobserved terms like error. Endogeneity implies that coefficients are biased and inconsistent if not properly estimated, typically requiring instrumental variables techniques [2]. We also did not consider alternative estimation methods (e.g., Bayesian methods) and alternative heterogeneity specifications (e.g., latent class models, a mixed logit model with joint parameter distributions, mixture models, and generalized logit models that account for scale and coefficient heterogeneity [69]).

Our study uses random utility discrete choice models that treat consumers as observant rational utility maximizers with consistent preferences. While this is a popular approach to modeling consumer choice, important criticisms exist. For instance, preferences can evolve over time [25], changing with cultural symbolism [70] and/or social interactions [71]. The theory of construction of preference adapted to design by MacDonald et al. [14] suggests that consumers' preferences for attributes do not exist a priori but are rather evaluated on a case-by-case basis [14]. Morrow and MacDonald [10] suggest that vehicle choice behavior may be better represented by a "consider-then-choose" model [72] where consumers first screen out most alternatives using simple rules, subsequently maximizing utility over a smaller "consideration set" [73]. The potential value of this type of model is suggested here by the better performance of class-only models, a special case of the consider-then-choose model. More broadly, the Lucas critique warns against use of aggregated historical data to predict outcomes in counterfactual future scenarios [74].

6 Conclusions

While the topic of uncertainty associated with choice predictions is widely discussed in the design community (e.g., Refs. [3–5,12,14,28]), there is no current consensus as to what processes and measures best quantify model uncertainty. This gap motivated our first research question, Q1. We investigated several well-known measures of model performance evaluated on a prediction set. For the automotive case study examined, likelihood measures (and the rank-equivalent KL divergence measure) tend to identify the same top-ranked model as the penalized likelihood measures AIC and BIC do. While CDFET measures identify different top-ranked models, depending on the error tolerance selected, the resulting models share most covariates. Models that perform well on one measure tend to perform well on the other measures, and models that perform poorly on one measure also tend to perform poorly on the other measures. In other words, determination of the best models in our study did not depend strongly on potentially arbitrary selection of the measure used to evaluate predictive accuracy.

Overall our results confirm several intuitive features of this application: attribute-based models predict better than models with no information; models of a particular vehicle class typically make better predictions than models of the full market; including more covariates generally improves predictive accuracy; and better model fit correlates well with better predictive accuracy. The match between fit and predictive accuracy, suggesting no major overfitting issues, is particularly encouraging, since the modeler has access to choice data for estimation but not choice data in the counterfactual predictive context. These findings would have to be validated in other product domains on a case-by-case basis.

We also observe a number of less intuitive results that are relevant to design. First, the models we construct are fairly poor predictors of future shares. In our base scenario, our best predictive model has an average error of 0.24% (the average share of a vehicle design is 0.42%), which translates to an error of approximately 37,500 vehicles sold for the 2007 market. The limited predictive power of standard models on real data in a canonical product category suggests designers should apply discrete choice models cautiously, though predictions may be substantially better in domains with fewer unobserved attributes or with conjoint data (where all attributes are observed).

Second, we find that attribute-based models do not furnish the best predictions for short-run forecasts in stable market conditions; attribute-based models estimated on 2004–2006 data were outperformed in predicting 2007 shares by the "static" model that assumes no changes in shares. However, attribute-based models are superior to the static model when predicting new vehicles only, since the static model lacks information about new entrants. There are some intuitive reasons why the static model might perform better than attribute-based models for short term predictions of existing designs given relatively stable market conditions. First, the static model may implicitly capture effects related to omitted vehicle attributes neglected by attribute-based models. Second, the static model may predict well in the short-run simply because of "inertial" conditions specific to the automotive market, particularly multiperiod production schedules and inventory buildup that must ultimately be cleared over the short run using unobserved advertising and/or purchasing incentives.

Third, while including an appropriate set of product attributes as model covariates is important to improving predictive accuracy, the form those covariates take in the utility function is less important in this application. This implies that it may be less important to test many variations of utility function covariate form when constructing a model, but it also means that any design decisions (e.g., design optimization results) that are not robust to variation in utility function covariate form may not be justified given the near equivalence of alternative covariate form in fit and prediction error with market data. If different utility specifications lead to different design decisions but the data cannot discern which form best represents choices, then design decisions cannot be reliably based on any single specification.

Finally, we observe that some of the models with the best predictive accuracy have coefficients with unexpected signs—likely biased due to correlation with unobserved attributes. Despite good prediction accuracy in existing markets, where attribute correlations are similar from year to year, these models may misguide design efforts if the designer makes changes that do not follow correlations in the marketplace. For example, the sign of the coefficient for the gallons per mile (gpm) attribute of the best predictive logit model is negative,[5] suggesting that consumers prefer lower fuel economy, all other attributes being equal. In fact, consumers may purchase vehicles with lower fuel economy because of other features of those vehicles unobserved by the modeler (e.g., size, performance, or styling attributes not captured in the model). The model predicts well if the new market retains such correlations, but a designer who lowers fuel economy alone is not likely to obtain the outcome predicted by the model. Thus, accuracy of predictions in existing markets is not a sufficient condition for use in design.

To verify that our results are not specific to the 2004–2006 timeframe, we conducted a similar analysis with estimation data from years 1971–1973 and 1981–1983 with prediction data from the respective one and four year forward markets. We find that our conclusions are generally robust to alternate timeframes: our accuracy measures are concordant; the best models exhibit substantial prediction error stemming from limited model fit; the static model outperforms the attribute-based models when predicting the full market one year forward but attribute-based models can predict better for four year forward forecasts or new vehicle designs; share predictions are sensitive to the presence of utility covariates but less sensitive to covariate form; nested and mixed logit specifications do not produce significantly more accurate forecasts; and the 1971–1973 models with best predictions do not necessarily have expected coefficient signs (though 1981–1983 models do). See Appendix H of the supplemental material for additional detail.

References

[1] Hazelrigg, G. A., 1998, "A Framework for Decision-Based Engineering Design," ASME J. Mech. Des., **120**(4), pp. 653–658.
[2] Train, K. E., 2009, *Discrete Choice Models With Simulation*, Cambridge University, New York.
[3] Wassenaar, H. J., and Chen, W., 2003, "An Approach to Decision-Based Design With Discrete Choice Analysis for Demand Modeling," ASME J. Mech. Des., **125**(3), pp. 490–497.
[4] Wassenaar, H. J., Chen, W., Cheng, J., and Sudjianto, A., 2005, "Enhancing Discrete Choice Demand Modeling for Decision-Based Design," ASME J. Mech. Des., **127**(4), pp. 514–523.
[5] Besharati, B., Luo, L., Azarm, S., and Kannan, P. K., 2006, "Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach," ASME J. Mech. Des., **128**(4), pp. 884–892.
[6] Williams, N., Azarm, S., and Kannan, P. K., 2008, "Engineering Product Design Optimization for Retail Channel Acceptance," ASME J. Mech. Des., **130**(6), p. 061402.
[7] Shiau, C.-S. N., and Michalek, J. J., 2009, "Optimal Product Design Under Price Competition," ASME J. Mech. Des., **131**(7), p. 071003.
[8] Frischknecht, B. D., Whitefoot, K., and Papalambros, P. Y., 2010, "On the Suitability of Econometric Demand Models in Design for Market Systems," ASME J. Mech. Des., **132**(12), p. 121007.
[9] He, L., Chen, W., Hoyle, C., and Yannou, B., 2012, "Choice Modeling for Usage Context-Based Design," ASME J. Mech. Des., **134**(3), p. 031007.
[10] Morrow, W. R., Long, M., and MacDonald, E. F., 2012, "Models in Decision Based Design Optimization," ASME DETC2012-71176.
[11] Whitefoot, K. S., and Skerlos, S. J., 2012, "Design Incentives to Increase Vehicle Size Created From the U.S. Footprint-Based Fuel Economy Standards," Energy Policy, **41**(1), pp. 402–411.
[12] Resende, C. B., Grace Heckmann, C., and Michalek, J. J., 2012, "Robust Design for Profit Maximization With Aversion to Downside Risk From Parametric Uncertainty in Consumer Choice Models," ASME J. Mech. Des., **134**(10), p. 100901.
[13] Payne, J. W., 1976, "Task Complexity and Contingent Processing in Decision Making: An Information Search," Organ. Behav. Hum. Perform., **16**(2), pp. 366–387.
[14] MacDonald, E. F., Gonzalez, R., and Papalambros, P. Y., 2009, "Preference Inconsistency in Multidisciplinary Design Decision Making," ASME J. Mech. Des., **131**(3), p. 031009.

[5]We reject the null hypothesis that the coefficient is equal to zero at the $\alpha = 0.01$ level for a two-sided *t*-test.

[15] Morrow, W. R., Long, M., and MacDonald, E. F., 2014, "Market-System Design Optimization With Consider-Then-Choose Models," ASME J. Mech. Des., **136**(3). p. 031003.
[16] Knittel, C. R., and Metaxoglou, K., 2012, "Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective," Rev. Econ. Stat., **96**(1), pp. 34–59.
[17] Allcott, H., and Wozny, N., 2012, "Gasoline Prices, Fuel Economy, and the Energy Paradox," Rev. Econ. Stat. (to be published).
[18] Dubé, J.-P., Fox, J. T., and Su, C.-L., 2012, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," Econometrica, **80**(5), pp. 2231–2267.
[19] Beresteanu, A., and Li, S., 2011, "Gasoline Prices, Government Support, and the Demand for Hybrid Vehicles in the United States," Int. Econ. Rev. (Philadelphia), **52**(1), pp. 161–182.
[20] Copeland, A., Dunn, W., and Hall, G., 2011, "Inventories and the Automobile Market," RAND J. Econ., **42**(1), pp. 121–149.
[21] Hoyle, C., Chen, W., Wang, N., and Gomez-Levi, G., 2011, "Understanding and Modelling Heterogeneity of Human Preferences for Engineering Design," ASME J. Eng. Des., **22**(8), pp. 583–601.
[22] Li, S., Liu, Y., and Zhang, J., 2011, "Lose Some, Save Some: Obesity, Automobile Demand, and Gasoline Consumption in the United States," J. Environ. Econ. Manage., **61**(1), pp. 52–66.
[23] Musti, S., and Kockelman, K. M., 2011, "Evolution of the Household Vehicle Fleet: Anticipating Fleet Composition, PHEV Adoption and GHG Emissions in Austin, Texas," Transp. Res. Part A Policy Pract., **45**(8), pp. 707–720.
[24] Zhang, T., Gensler, S., and Garcia, R., 2011, "A Study of the Diffusion of Alternative Fuel Vehicles: An Agent-Based Modeling Approach," J. Prod. Innov. Manag., **28**(2), pp. 152–168.
[25] Axsen, J., Mountain, D. C., and Jaccard, M., 2009, "Combining Stated and Revealed Choice Research to Simulate the Neighbor Effect: The Case of Hybrid-Electric Vehicles," Resour. Energy Econ., **31**(3), pp. 221–238.
[26] Dagsvik, J. K., and Liu, G., 2009, "A Framework fo Analyzing Rank-Ordered Data With Application to Automobile Demand," Transp. Res. Part A Gen., **43**(1), pp. 1–12.
[27] Hoyle, C., Chen, W., Ankenman, B., and Wang, N., 2009, "Optimal Experimental Design of Human Appraisals for Modeling Consumer Preferences in Engineering Design," ASME J. Mech. Des., **131**(7), p. 071008.
[28] Shiau, C.-S. N., and Michalek, J. J., 2009, "Should Designers Worry About Market Systems?," ASME J. Mech. Des., **131**(1), p. 011011.
[29] Vance, C., and Mehlin, M., 2009, *Tax Policy and CO_2 Emissions—An Econometric Analysis of the German Automobile Market*, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen, Bochum, Dortmund, Duisburg, Essen, Germany.
[30] Dasgupta, S., Siddarth, S., and Silva-Risso, J., 2007, "To Lease or to Buy? A Structural Model of a Consumer's Vehicle and Contract Choice Decisions," J. Mark. Res., **44**(3), pp. 490–502.
[31] Train, K., and Winston, C., 2007, "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," Int. Econ. Rev. (Philadelphia), **48**(4), pp. 1469–1497.
[32] Kemper, L., 2006, "Multilevel Optimization for Enterprise-Driven Decision-Based Product Design," *Decision Making in Engineering*, L. Kemper, W. Chen, and L. Schmidt, eds., ASME Press, NY, pp. 203–214.
[33] Berry, S., Levinsohn, J., and Pakes, A., 2004, "Differentiated Products Demand Systems From a Combination of Micro and Macro Data: The New Car Market," J. Polit. Econ., **112**(1), pp. 68–105.
[34] Greene, D. L., Patterson, P. D., Singh, M., and Li, J., 2005, "Feebates, Rebates and Gas-Guzzler Taxes: A Study of Incentives for Increased Fuel Economy," Energy Policy, **33**(6), pp. 757–775.
[35] Michalek, J. J., Papalambros, P. Y., and Skerlos, S. J., 2004, "A Study of Fuel Efficiency and Emission Policy Impact on Optimal Vehicle Design Decisions," ASME J. Mech. Des., **126**(6), pp. 1062–1070.
[36] Petrin, A., 2002, "Quantifying the Benefits of New Products: The Case of the Minivan," J. Polit. Econ., **110**(4), pp. 705–729.
[37] Sudhir, K., 2001, "Competitive Pricing Behavior in the Auto Market: A Structural Analysis," Mark. Sci., **20**(3), pp. 42–60.
[38] Berry, S., Levinsohn, J., and Pakes, A., 1999, "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," Am. Econ. Rev., **89**(3), pp. 400–430.
[39] Brownstone, D., and Train, K., 1999, "Forecasting New Product Penetration With Flexible Substitution Patterns," J. Econ., **89**(1–2), pp. 109–129.
[40] Dagsvik, J. K., Wennemo, T., Wetterwald, D. G., and Aaberge, R., 2002, "Potential Demand for Alternative Fuel Vehicles," Transp. Res. Part B: Methodol., **36**(4), pp. 361–384.
[41] McCarthy, P. S., 1996, "Market Price and Income Elasticities of New Vehicle Demands," Rev. Econ. Stat., **78**(3), pp. 543–547.
[42] Berry, S., Levinsohn, J., and Pakes, A., 1995, "Automobile Prices in Market Equilibrium," Econometrica, **63**(4), pp. 841–890.
[43] Boyd, J. H., and Mellman, R. E., 1980, "The Effect of Fuel Economy Standards on the US Automotive Market: A Hedonic Demand Analysis," Transp. Res. Part A Gen., **14**, pp. 367–378.
[44] Lave, C., and Train, K., 1979, "A Disaggregate Model of Auto-Type Choice," Transp. Res. Part A Gen., **13**(1), pp. 1–9.
[45] Feit, E. M., Beltramo, M. A., and Feinberg, F. M., 2010, "Reality Check: Combining Choice Experiments with Market Data to Estimate the Importance of Product Attributes," Manage. Sci., **56**(5), pp. 785–800.
[46] Bunch, D., Greene, D., Lipman, T., and Shaheen, S., 2011, *Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California*, University of California, Davis, Institute of Transportation Studies, Davis, CA.

[47] U.S. Energy Information Administration, 2011, *Annual Energy Outlook 2011 With Projections to 2035*, U.S. Energy Information Administration, Washington, DC.
[48] Gigerenzer, G., and Brighton, H., 2009, "Homo Heuristicus: Why Biased Minds Make Better Inferences," Top. Cogn. Sci., **1**(1), pp. 107–143.
[49] Train, K., 1979, "Consumers' Responses to Fuel-Efficient Vehicles: A Critical Review of Econometric Studies," Transportation (Amsterdam), **8**(3), pp. 237–258.
[50] Chen, C., 2001, "Design for the Environment: A Quality-Based Model for Green Product Development," Manage. Sci., **47**(2), pp. 250–263.
[51] Shiau, C.-S. N., Michalek, J. J., and Hendrickson, C. T., 2009, "A Structural Analysis of Vehicle Design Responses to Corporate Average Fuel Economy Policy," Transp. Res. Part A: Policy Pract., **43**(9–10), pp. 814–828.
[52] Nevo A., 2000, "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," J. Econ. Manage. Strategies, **9**(4), pp. 513–548.
[53] Dubé, J., Fox, J., and Su, C., 2012, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," Econometrica **80**(5), pp. 2231–2267.
[54] Skrainka, B., and Judd, K., 2011, "High Performance Quadrature Rules: How Numerical Integration Affects a Popular Model of Product Differentiation," (published online). Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1870703
[55] Bunch, D. S., Bradley, M., and Golob, T. F., 1993, "Demand for Clean Fuel Vehicles in California A Discrete Choice Stated Preference Pilot Project," Transp. Res. Part A: Policy Pract., **27A**(3), pp. 237–253.
[56] Brownstone, D., Bunch, D. S., and Train, K., 2000, "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-Fuel Vehicles," Transp. Res. Part B: Methodol., **34**(5), pp. 315–338.
[57] McFadden, D., and Train, K., 2000, "Mixed MNL Models for Discrete Response," J. Appl. Econ., **15**(5), pp. 447–470.
[58] Selby, B., and Kockelman, K., 2012, "Microsimulating Automobile Markets: Evolution of Vehicle Holdings and Vehicle Pricing Dynamics," Transp. Res. Forum, **51**(2), pp. 83–96.
[59] Ward's Automotive Group, 2008, "Ward's Yearly Vehicle Specifications." Available at: http://wardsauto.com/data-center
[60] Polk, 2008, "IHS Polk Insight Yearly Volume Data." Available at: http://blog.polk.com/blog/new-vehicle-sales
[61] U.S. Energy Information Administration, 2013, "U.S. All Grades All Formulations Retail Gasoline Prices" (accessed: Aug. 1, 2013) http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPM0_PTE_NUS_DPG&f=A
[62] "Annual Average Indexex," Bur. Labor Statitstics Data Online, (accessed: Aug. 1, 2013), http://www.bls.gov/cpi/cpi_dr.htm#2007
[63] Wooldridge, J. M., 2006, *Introductory Econometrics: A Modern Approach*, Thomas Higher Education, Mason, OH.
[64] Kullback, S., and Leibler, R. A., 1951, "On Information and Sufficiency," Ann. Math. Stat., **22**(1), pp. 79–86.
[65] Akaike, H., 1973, "Information Theory and an Extension of the Maximum Likelihood Principle," Proceedings of the Second International Symposium on Information Theory, B. N. Petrov and F. Csaki, eds., Akademiai Kiado, Budapest, Hungary, pp. 267–281.
[66] Schwarz, G., 1978, "Estimating the Dimension of a Model," Ann. Stat., **6**(2), pp. 461–464.
[67] Burnham, K. P., and Anderson, D. R., 2002, *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, Springer, New York.
[68] Claeskens, G., and Hjort, N., 2008, *Model Selection and Model Averaging*, Cambridge University Press, NY.
[69] Fiebig, D. G., Keane, M. P., Louviere, J., and Wasi, N., 2009, "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity," Mark. Sci., **29**(3), pp. 393–421.
[70] Heffner, R. R., Kurani, K. S., and Turrentine, T. S., 2007, "Symbolism in California's Early Market for Hybrid Electric Vehicles," Transp. Res. Part D: Transp. Environ., **12**(6), pp. 396–413.
[71] Axsen, J., and Kurani, K. S., 2011, "Interpersonal Influence in the Early Plug-in Hybrid Market: Observing Social Interactions With an Exploratory Multi-Method Approach," Transp. Res. Part D: Transp. Environ., **16**(2), pp. 150–159.
[72] Hauser, J., and Wernerfelt, B., 1990, "An Evaluation Cost Model of Consideration Sets," J. Consum. Res., **16**(4), pp. 393–408.
[73] Gilbride, T. J., and Allenby, G. M., 2004, "A Choice Model With Conjunctive, Disjunctive, and Compensatory Screening Rules," Mark. Sci., **23**(3), pp. 391–406.
[74] Lucas, R. E., 1976, "Econometric Policy Evaluation: A Critique," Carnegie-Rochester Conference Series on Public Policy, **1**, pp. 19–46.

Environmental Research Letters

OPEN ACCESS

RECEIVED
24 August 2015

REVISED
24 December 2015

ACCEPTED FOR PUBLICATION
5 January 2016

PUBLISHED
9 February 2016

LETTER

Consequential life cycle air emissions externalities for plug-in electric vehicles in the PJM interconnection

Allison Weis[1], **Paulina Jaramillo**[1] **and Jeremy Michalek**[1,2]

[1] Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Ave, PA 15213, USA
[2] Mechanical Engineering, Carnegie Mellon University, 5000 Forbes Ave, PA 15213, USA

E-mail: paulina@cmu.edu

Keywords: electric vehicles, LCA, social costs

Supplementary material for this article is available online

Abstract

We perform a consequential life cycle analysis of plug-in electric vehicles (PEVs), hybrid electric vehicles (HEVs), and conventional gasoline vehicles in the PJM interconnection using a detailed, normative optimization model of the PJM electricity grid that captures the change in power plant operations and related emissions due to vehicle charging. We estimate and monetize the resulting human health and environmental damages from life cycle air emissions for each vehicle technology. We model PJM using the most recent data available (2010) as well as projections of the PJM grid in 2018 and a hypothetical scenario with increased wind penetration. We assess a range of sensitivity cases to verify the robustness of our results. We find that PEVs have higher life cycle air emissions damages than gasoline HEVs in the recent grid scenario, which has a high percentage of coal generation on the margin. In particular, battery electric vehicles with large battery capacity can produce two to three times as much air emissions damage as gasoline HEVs, depending on charge timing. In our future 2018 grid scenarios that account for predicted coal plant retirements, PEVs would produce air emissions damages comparable to or slightly lower than HEVs.

1. Introduction

Plug-in electric vehicle (PEV) technologies, including plug-in hybrid electric vehicles (PHEVs) and battery electric vehicles (BEVs), have the potential to reduce environmental impacts from the transportation system by reducing or eliminating tailpipe emissions. However, the emissions associated with producing PEVs and generating the electricity to charge PEVs affect whether these vehicles have higher or lower life cycle environmental and health impacts compared to efficient gasoline vehicles (Michalek *et al* [1], Tessum *et al* [2]). Evaluating the sustainability of different transportation choices thus requires both a consideration of the full life cycle of the technology as well as an analysis of the impacts of the technology choice [3]. While there has been significant research to understand the life cycle environmental impacts of PEVs, most of this prior work has followed an attributional life cycle assessment (LCA) approach to answer the question '*what air emissions are PEV charging responsible for*'? Such approaches have assumed that PEV charging produces emissions proportional to the average emissions rate for electricity generation in the political boundary (country, state, etc) or grid region (NERC region, eGRID sub-region, interconnect, balancing area, etc) where the vehicle is charged. Alternatively, some studies have used hypothetical emissions factors to evaluate a broader set of scenarios of the effects of PEV charging [1, 2, 4, 5]. The results from previous work thus vary depending on the researcher's value judgment related to the emissions that a PEV should be responsible for when charging in a particular location.

In contrast, a consequential approach answers '*what are the air emissions implications of PEV adoption in a region*' by assessing how grid operations change in response to new charging demand. These consequential effects of PEV charging on grid emissions have been examined using empirical 'top-down' methods and normative 'bottom-up' methods [6]. Top-down empirical models of the power system, like those developed by Siler-Evans *et al* [7] and Graff Zivin *et al* [8], use regressions on historical data to estimate marginal emissions rates. Such analysis is grounded in the actual operations of the power system. However, the approach is limited to historical scenarios and is only appropriate for the analysis of small changes in generation or load. It also suffers from error in counterfactual analysis because correlations in past data do not necessarily imply causality. The gap between correlation and causality is particularly evident for the dispatch of hydroelectric plants, which may change generation timing in response to new load but typically will not change total energy generated in response to new load. Alternatively, bottom-up normative models of the power system, such as those used by Sioshansi *et al* [9], Peterson *et al* [10], Choi *et al* [11] and Weis *et al* [12], use optimization models to estimate how a power system should operate to minimize costs subject to a variety of constraints. These models can assess changes of grid operation in response to new PEV load. Such analysis can model future power plant scenarios and large load changes. While this approach has limited scalability for modeling large systems, and it is typically not possible to model all possible considerations that affect grid operations in practice, there is growing interest in using bottom-up models of the power system in consequential LCA.

Table 1 summarizes prior consequential LCA studies of PEV air emissions. Tamayao *et al* [6] apply two different top-down regression models [4, 5] to assess consequential greenhouse gas (GHG) emissions in the US, while Ma *et al* [13] perform their own regression on the operation of the UK grid. Sandy [14] and Onat *et al* [15] average results from ORNL's analysis of marginal emissions for electric vehicles using a dispatch model and different vehicle charging patterns [16]. Finally, Choi *et al* [11] construct a bottom-up capacity expansion and unit commitment model to assess consequential GHG emissions. In this paper, we adopt the bottom-up normative approach to model the power system under different scenarios in the PJM system in order to inform a consequential LCA of PEVs. Unlike previous work, our analysis includes the valuation of social damages associated with emissions of criteria air pollutants as well as GHG emissions. PJM (an independent system operator in Pennsylvania, New Jersey, Maryland, Ohio, and several other states) is an interesting power system to examine, as it is the largest independent system operator in the United States by population and has a large installed coal capacity. The supplemental information (SI) includes an expanded comparison of this study with previous life cycle studies of PEVs in the United States.

2. Methods

2.1. Life cycle boundary

We estimate the life cycle emissions of CO_2, CO, SO_2, $PM_{2.5}$, NO_x, and VOCs for conventional, hybrid, and PEVs, including the emissions from vehicle manufacturing, fuel production, and use. The analysis focuses on the region covered by the PJM interconnection, which has the largest electricity market in the US and serves 13 states in the Mid-Atlantic region [17]. For CVs and HEVs, the emissions from fuel production and use include upstream emissions from petroleum drilling and refining as well as the tailpipe emissions during vehicle operations. For BEVs, fuel production emissions include power plant emissions and upstream emissions from coal and natural gas production, but these vehicles do not have tailpipe emissions. Finally, PHEVs have emissions associated with both gasoline and electricity. Figure 1 shows the scope of the life cycle inventory. We do not consider end of life emissions. Further, we assume a total vehicle life of 160 000 miles for all vehicle types, following Michalek *et al* [1].

2.2. Vehicle and power grid scenarios

For this analysis we rely on a unit commitment and economic dispatch (UCED) model previously developed and used to evaluate the costs and benefits of controlled charging of electric vehicles in the PJM system [12, 18]. This model uses a mixed integer linear optimization program to minimize the costs of operating the power system to meet load given constraints, including power plant operating constraints (minimum and maximum load, ramping rates, and minimum up and down time), as well as transmission and operating reserves constraints. Furthermore, the model incorporates vehicle charging by including charging requirements and battery constraints.

In order to account for changes in the composition of the PJM power plant fleet that may result from environmental regulations, renewable energy mandates, and changes in energy prices, we develop three scenarios of the power system. For the first scenario, we use the Environmental Protection Agency (EPA)'s NEEDS database in order to represent the recent PJM system [19]. Our second scenario, meant to represent a future grid, includes retirement of power plants predicted by the EPA [20] and a 3% wind penetration, as described in Weis *et al* [12]. Finally, in the third scenario we include the power plant retirements EPA predicts, and we add 20% wind. The additional wind generation for the future grid comes from NREL's Eastern Wind Integration and Transmission Study dataset [21]. For each scenario, we add wind sites in order of capacity factor to reach the required wind penetration level (3% for the second scenario and 20% for the third scenario).

Table 1. Summary of literature estimating consequential life cycle air emissions from PEVs.

Study	Location	Air emissions	Model type	Model detail
Tamayao *et al* [6]	US	GHG	Empirical	Applies regression results from [4, 5]
Choi *et al* [11]	Eastern Interconnect	GHG	Normative	Capacity expansion and unit commitment
Ma *et al* [13]	CA, UK	GHG	Empirical	Regression on plant operation to find average marginal emission factor
Sandy [14]	US	GHG	Normative	Averages resulting emissions from power plant dispatch in [16]
Onat *et al* [15]	US	GHG	Normative	Averages resulting emissions from power plant dispatch in [16]
This study	US PJM	GHG, SO_2, NO_x, $PM_{2.5}$, CO, VOC	Normative	Optimal dispatch

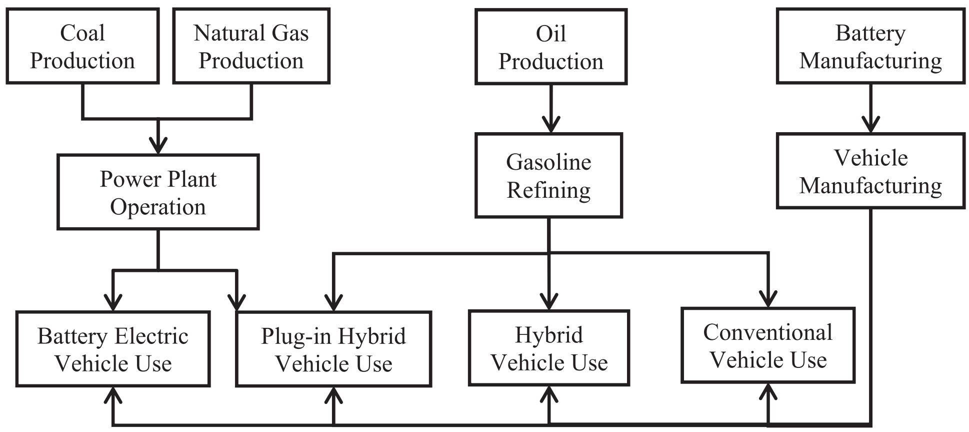


Figure 1. Life cycle inventory for plug-in electric, hybrid electric, and conventional vehicles.

We also account for differences in vehicle technology using Argonne National Laboratory's 2013 GREET 1 and 2 models [22, 23]. The base case PEV for the recent grid scenario is a 2010 PHEV-35, sized to represent the Chevy Volt. We also include a smaller battery size based on the Toyota Plug-in Prius and a larger battery size based on the Tesla Model S. The efficiency of the long-range BEV is that of the 2012 Tesla Model S, as measured by the EPA [24] (we examine more efficient BEVs in sensitivity analysis). For the future grid scenarios we use GREET's 2015 PHEV-35 vehicle specifications. Table 2 summarizes the differences across scenarios.

2.3. Life cycle inventory

We determine the life cycle emissions for each stage shown in figure 1 for each vehicle type based on a 160 000 mile vehicle life. Table 3 provides a summary of the data used for each stage of the life cycle.

2.3.1. Emissions from power plant operations

In order to estimate the change in emissions from the power system resulting from vehicle charging, we solve the UCED model both with and without electric vehicles and compare the difference in the operating schedules of the power plants. We add the charging load to the existing non-vehicle electricity load by assuming electric vehicles make up 10% of the vehicle fleet in PJM. Furthermore, we assume that electric vehicles are distributed throughout the PJM system proportionally to population (we test alternative assumptions in sensitivity analysis).

Using data from the National Household Travel Survey (NHTS) [25] we estimate the vehicle charging load, following Weis *et al* [12, 26]. The NHTS data provide the distance driven during each trip throughout the day surveyed as well as the time of each trip for approximately 100 000 passenger cars across the US. We use the distance driven in a day, the vehicle efficiency, and the electric range of the vehicle to calculate both the distance driven in charge-depleting versus charge-sustaining mode and the total charging load per day. We assume that all PHEVs drive as far as possible in charge-depleting mode before switching to charge-sustaining mode. We also assume that vehicle charging occurs at home after the last trip of the day and that each vehicle is fully charged by the first trip of the next day. For each fleet/vehicle type scenario, we calculate the hourly charging load for a scenario where charging begins at full power following the last trip of the day ('uncontrolled charging') and a scenario where the utility dynamically controls the rate at which each vehicle is charged in order to minimize generation cost ('controlled charging'). Controlled charging of electric vehicles may provide operational benefits to the grid, but it also changes the emissions associated with electric vehicle charging [12], so it represents an additional scenario for analysis. For the controlled charging scenarios, we use 20 representative vehicle profiles to model the load from the electric vehicles as described in Weis *et al* [12] and assume that the vehicle is constrained to be fully charged by the first trip of the following day while the charging rate is limited to Level 2 charging (7.2 kW).

2.3.2. Tailpipe emissions

The distance driven using the gasoline engine (all times for the CV and HEV and during charge-sustaining mode for PHEVs) and the vehicle's emission rate determine its tailpipe emissions. We assume that each vehicle travels 160 000 miles over its lifetime. Table 4 shows tailpipe emission rates for GHG, CO, SO_2, $PM_{2.5}$, NO_x, and VOCs from the GREET 1 model [22]. While the two PHEVs burn some gasoline in charge-depleting mode, the BEV-265 operates without any tailpipe emissions.

2.3.3. Upstream emissions

Our life cycle emissions include the emissions from the production, processing, and delivery of fossil fuels (either to fuel the vehicles or to generate electricity), as well as the emissions from vehicle and battery

Table 2. Summary of scenarios.

Scenario	Power system data	Conventional vehicle	Hybrid vehicle	Plug-in electric vehicles
Recent (2010)	2010 PJM	2010 GREET ICEV	2010 GREET HEV	2010 GREET PHEV-10 (Plug-in Prius-sized) 2010 GREET PHEV-35 (Volt-sized) 2012 BEV-265 (Tesla-sized)
Future (~2018)	EPA forecasted 2018 PJM with 3% wind penetration	2015 GREET ICEV	2015 GREET HEV	2015 GREET PHEV-35 (Volt-sized)
Hypothetical high wind future	EPA forecasted 2018 PJM with 20% wind penetration	2015 GREET ICEV	2015 GREET HEV	2015 GREET PHEV-35 (Volt-sized)

Table 3. Data for the life cycle emissions for each stage.

Stage	Emission rate	Source	Other assumptions	Source
Power plant operation	Short Ton/year	Unit commitment model	Driving patterns and vehicle efficiency	NHTS GREET 1, fueleconomy.gov
Tailpipe	lb/mile	GREET 1	Driving patterns	NHTS
Vehicle manufacturing	Short Ton/lifetime	GREET 2		
Battery manufacturing	Short Ton/lifetime	GREET 2		
Oil production	Short Ton/mile	GREET 1		
Gasoline refining	Short Ton/mile	GREET 1		
Coal production	Short Ton/MWh	GREET 1	MWh produced using coal	Unit commitment model
Natural gas production	Short Ton/MWh	GREET 1	MWh produced using natural gas	Unit commitment model

Table 4. Tailpipe emissions in pounds per 1000 miles from GREET 1. The recent grid scenario relies on the data for the 2010 vehicles, while the future scenarios are based on the characteristics of the 2015 vehicles. CS: charge-sustaining mode. CD: charge-depleting mode.

Vehicle	CO_2-eq	VOC	CO	NO_x	$PM_{2.5}$	SO_2
2010 CV	772	0.375	6.393	0.265	0.026	0.011
2010 HEV	551	0.265	6.393	0.220	0.026	0.008
2010 PHEV-10 CS (Plug-in Prius)	529	0.265	6.393	0.220	0.026	0.008
2010 PHEV-10 CD (Plug-in Prius)	265	0.088	2.205	0.079	0.006	0.004
2010 PHEV-35 CS (Volt)	683	0.265	6.393	0.220	0.026	0.010
2010 PHEV-35 CD (Volt)	44.1	0.022	0.375	0.013	0.001	0.001
2010 BEV-265 (Tesla Model S)	0.00	0.00	0.00	0.00	0.00	0.00
2015 CV	705	0.375	6.393	0.265	0.026	0.011
2015 HEV	507	0.265	6.393	0.220	0.026	0.007
2015 PHEV CS (Volt)	573	0.265	6.393	0.220	0.026	0.009
2015 PHEV CD (Volt)	41.9	0.016	0.397	0.014	0.001	0.001

manufacturing. Argonne National Laboratory's GREET 1 model [22] provides the emission rates for oil drilling and refining, which are the basis of the upstream emissions estimate for gasoline. Upstream emissions for vehicle and battery manufacturing come from the GREET 2 model [23].

To account for the emissions from the production, processing, and delivery of coal and natural gas for power plants used to charge the electric vehicles, we rely on the GREET 1 model, which provides these emissions on a per MWh basis (as shown in the supporting information). We then apply these emissions factors to the amount of electricity generated from coal and natural gas plants in the UCED model results. We include only the upstream emissions from coal- and natural gas-based electricity as these fuels account for the majority of generation response to additional vehicle charging load [12]. In the high wind scenario, wind power also contributes to vehicle charging, but wind requires no fuel and thus has no upstream emissions (our model does not include the emissions from building physical infrastructure).

2.4. Life cycle damages

Air emissions cause environmental degradation and affect human health. For CO_2, we use estimates of the social cost of carbon that the US EPA uses for regulatory impact assessment [27]. EPA reports this social cost of carbon for three different discount rates: 2.5%, 3%, and 5%. We use the 3% discount rate average value for 2010 as our base value for all scenarios.

In order to estimate the damages from SO_2, $PM_{2.5}$, NO_x, and VOCs we use the values from the AP2 model [28], which estimates the marginal health and environmental damages for emissions of each criteria air pollutant in each county in the United States. This model has many uncertain parameters, including the value of statistical life, which is used to monetize morbidity and mortality from air pollution. The AP2 model includes results from a Monte Carlo analysis of the damages in each county for the baseline year (2005). As a base case, we assume that these 2005 marginal damages per unit of emission in each location apply also to the recent grid and future grid scenarios. We use the distribution of the results from the Monte Carlo analysis to characterize the uncertainty within the AP2 model. Damages from vehicle charging are based on the change in the annual generation and emissions from each power plant in the UCED model that results from increased charging load. Since the AP2 values are specific to individual counties where emissions take place, we also need to incorporate the location of the vehicle tailpipe emissions. To do so, we allocate vehicles to counties within PJM proportionally to population. We further assume that each vehicle is driven within its respective county.

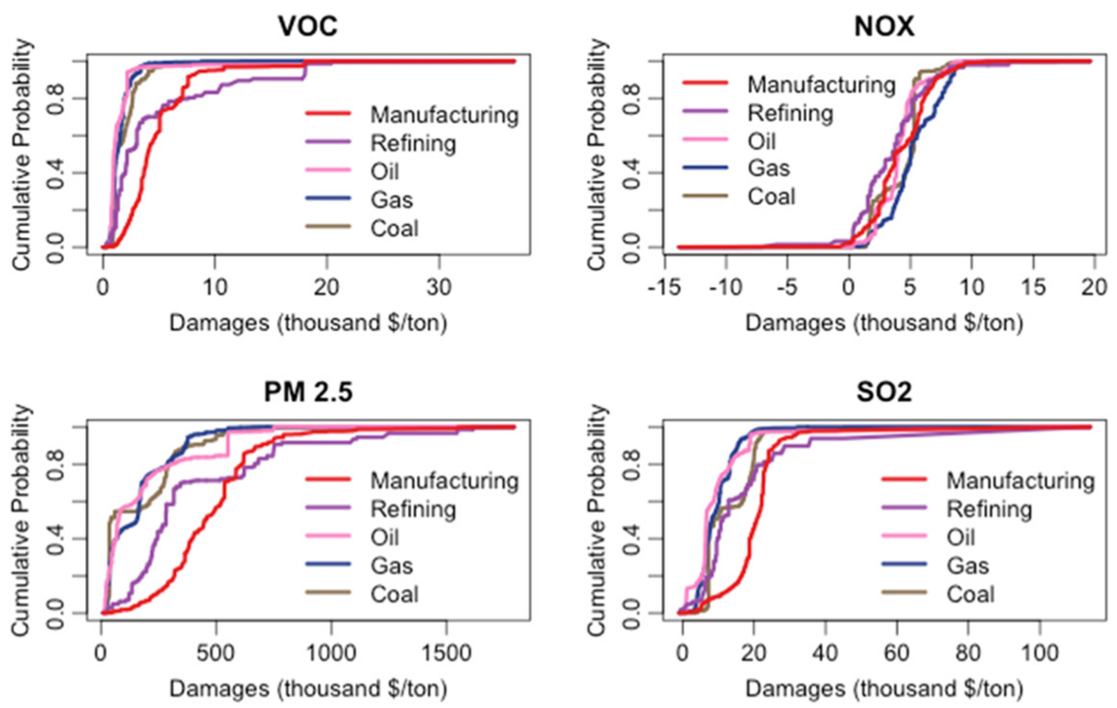


Figure 2. Cumulative probability distribution of damages per ton of pollutant released for assumed locations of upstream production emissions by pollutant type (using AP2 data for 2005).

We allocate the emissions from vehicle and battery manufacturing to US counties identified by the US census as having automobile and parts manufacturing activity, weighted by the number of automotive manufacturing workers, following Michalek *et al* [1]. We also allocate coal, oil, and natural gas upstream emissions to US counties where mines, oil and gas fields, and refineries are located, weighted by the production in each county. Figure 2 shows the resulting cumulative probability distribution of damages from manufacturing, coal, oil and gas production, and oil refining. The damage calculations assume that all emissions and damages occur in the United States, though in practice some of these processes occur outside of US borders. Marginal emissions in other countries could incur higher or lower damages than in the US, depending on the existing pollutant concentrations and populations in those areas. At this time, however, we are unable to include these damages. We allocate the damages for all life cycle stages except for vehicle and battery manufacturing across the years of the vehicle lifetime, ignoring changes in the electricity grid over the vehicle life. We then used a 3% discount rate to find the present value of these damages, consistent with the calculation of CO_2 damages.

3. Results

3.1. Life cycle emissions

Figure 3 shows the breakdown of estimated emissions by life cycle stage for each scenario. In the recent grid scenario, the PEVs have higher GHG, SO_2, NO_x, and $PM_{2.5}$ emissions and lower CO and VOC emissions than the HEV. Compared to the CV, PEVs have higher SO_2 emissions and lower CO and VOC emissions, while GHG, NO_x and $PM_{2.5}$ emissions may be higher or lower, depending on the PEV characteristics and the charging scenario. Controlled charging of PEVs increases emissions of GHGs, SO_2, NO_x, and $PM_{2.5}$ while reducing emissions of VOCs compared to uncontrolled charging, due in part to the increased use of coal-fired power plants available at night. Use of coal generation increases with controlled charging in all scenarios because these plants have lower marginal cost than gas plants and have more excess capacity available in off-peak hours than during uncontrolled charging ours (the hours immediately following the vehicle's first home arrival at the end of the day). For a detailed description of the analysis and assumptions regarding the response of power plant dispatch to additional electric vehicle charging load see [12].

In the future grid scenarios, compared to the HEV, the PHEV-35 has higher SO_2 emissions and lower $PM_{2.5}$, VOC, and CO emissions, while GHG and NO_x emissions may be higher or lower, depending on the charging scenario and wind power scenario. Compared to the CV, the PHEV-35 has higher SO_2 and lower GHG, NO_x, $PM_{2.5}$, VOC, and CO emissions. Controlled charging of the PHEV-35 increases SO_2 and NO_x while decreasing VOC emissions, but the effect on GHG and $PM_{2.5}$ emissions depends on the presence of wind, and the effect on CO emissions is negligible.

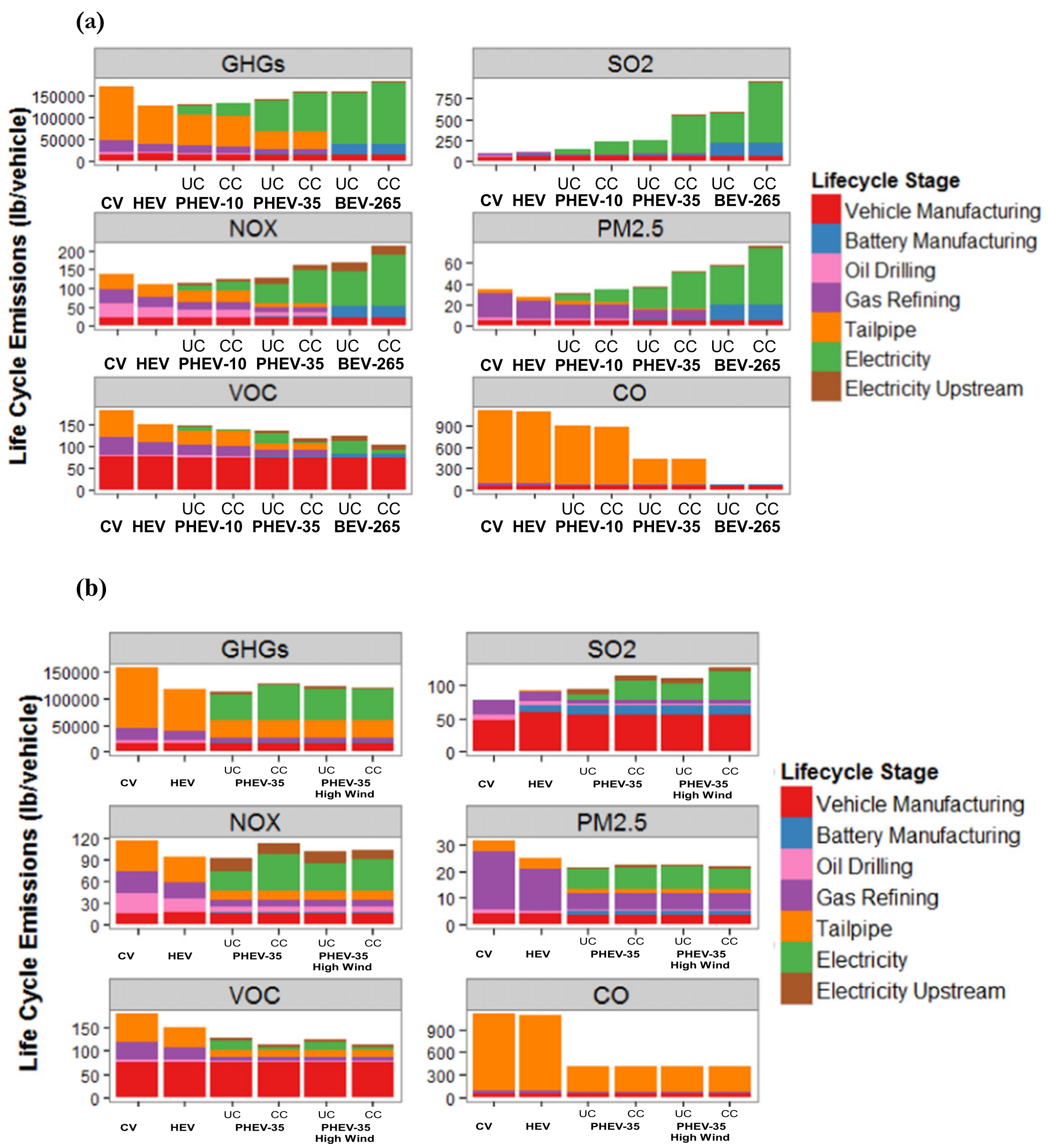


Figure 3. Life cycle emissions by pollutant and life cycle stage for each vehicle type in the recent (a) and future (b) PJM grid. UC stands for uncontrolled charging and CC stands for controlled charging.

3.2. Life cycle damages

3.2.1. Expected values

PEVs have higher expected life cycle damages than hybrid vehicles in the recent PJM scenario in all cases examined, as shown in figure 4. Their expected damages are also higher than those of conventional vehicles, except for the case of the PHEV-10 with uncontrolled charging. Long-range BEVs cause two to three times as much air emissions damage as HEVs. The electricity generation damages come largely as a result of the SO_2 emissions from the coal plants used to charge the vehicles in off-peak hours. Controlled charging increases life cycle damages relative to uncontrolled charging because of the increases in emissions associated with higher levels of coal generation. Uncertainty is not presented here because common sources of uncertainty create correlated uncertainty across scenarios, so error bars could be misleading in comparing across cases. Instead, section 3.2.2. includes an analysis of uncertainty and robustness in these results.

In the future scenarios, shown in figure 5, the PHEV-35 is able to reduce life cycle damages by a few hundred dollars over its lifetime compared to the CV and the HEV. Again, PEVs tend to produce larger damages under controlled charging than under uncontrolled charging, but in the future scenarios the PHEV-35 provides benefits compared to the CV and the HEV regardless of the charging scheme. The high-wind future scenarios do not necessarily imply lower consequential damages than the low-wind future scenarios for PEVs because added wind displaces fossil fuel plants, which can increase the availability of coal on the margin when PEVs charge.

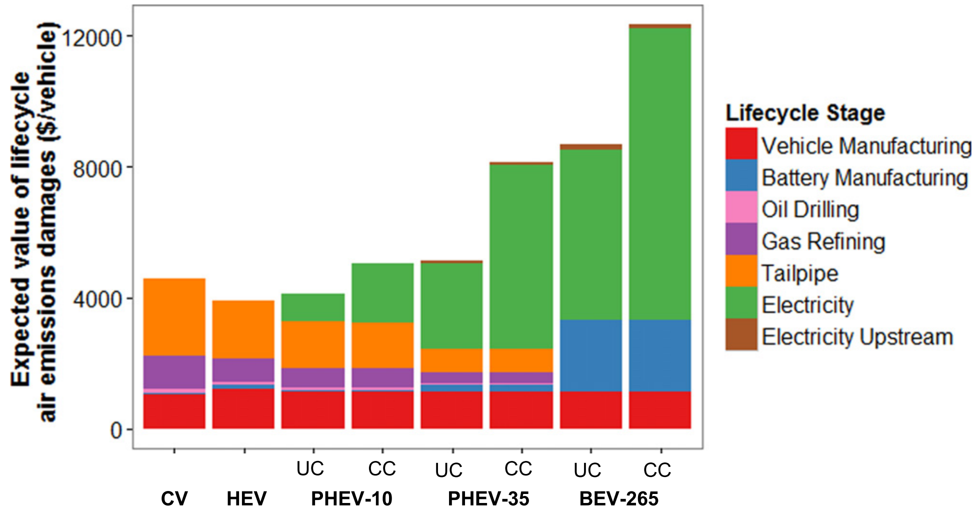


Figure 4. Expected value of life cycle air emission damages for each vehicle type in the recent PJM grid. UC stands for uncontrolled charging and CC stands for controlled charging for the electric vehicles.

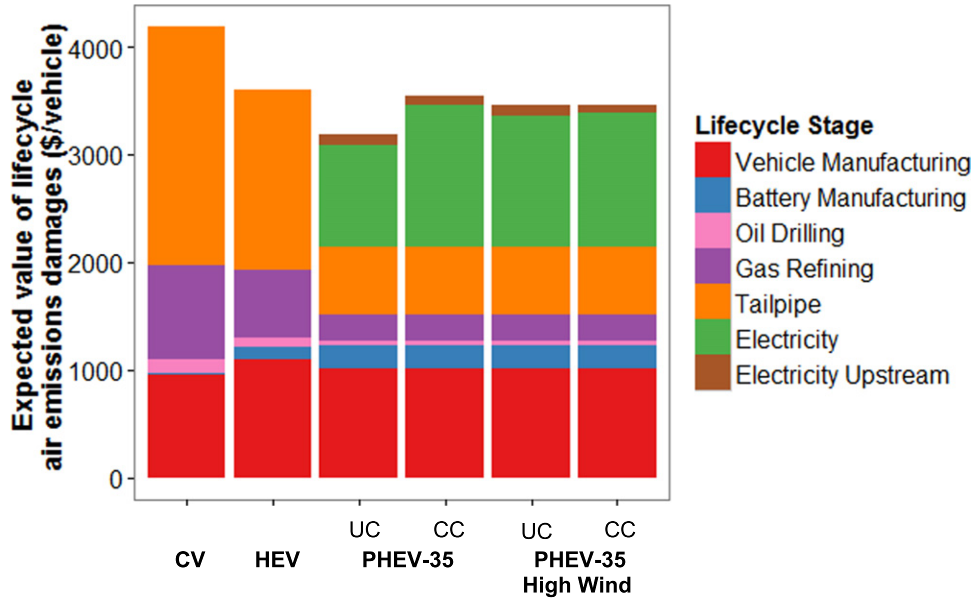


Figure 5. Expected value of life cycle damages in the future PJM grid. The high wind scenario has 20% of demand met by wind. UC stands for uncontrolled charging and CC stands for controlled charging.

3.2.2. Uncertainty and robustness

Qualitatively, our key findings are that (1) the PEVs cause more damage than the CV and HEV in the recent PJM grid, (2) the PHEV-35 causes less damage than the CV in the future PJM grid scenarios, but these damages are not much lower than those for the HEV, and (3) controlled charging tends to increase damages compared to uncontrolled charging, though the PEVs provide benefits compared to the CV and HEV in the future grid scenarios regardless of charging scheme. To characterize uncertainty and robustness of these findings, we use the Monte Carlo analysis results from the AP2 model and assess the probability that each vehicle technology has higher life cycle air emissions damages than the HEV. Table 5 reveals that the conclusions above are robust, especially in the recent grid. The uncertainty in the AP2 models does not significantly affect the comparison across vehicle types because most of this uncertainty is a result of uncertainty in the value of a statistical life, which affects damage estimates across all cases. As a result, this uncertainty typically changes only the magnitude of the difference between hybrids and other vehicles, not the sign.

The results in table 5 incorporate only the uncertainty quantified in the AP2 model due to uncertainty in input parameters, such as the value of a statistical life. We examine the effect of some of the other important input parameters and assumptions through sensitivity cases summarized in table 6

Table 5. Robustness of results for the damage difference between hybrid vehicles and each other vehicle type. CV = conventional vehicle.

Scenario	Charging	Probability damages are larger than for HEV's	Mean change in life cycle damages compared to HEV's
CV—recent grid		100%	\$650
PHEV-10—recent grid	Uncontrolled	81%	\$210
	Controlled	98%	\$1100
PHEV-35—recent grid	Uncontrolled	95%	\$1200
	Controlled	99%	\$4200
BEV-265—recent grid	Uncontrolled	99%	\$4800
	Controlled	99%	\$8400
CV—future grid		99%	\$580
PHEV-35—future grid	Uncontrolled	4%	−\$420
	Controlled	28%	−\$60
PHEV-35—future grid with high wind	Uncontrolled	17%	−\$150
	Controlled	18%	−\$150

Table 6. Summary of findings from sensitivity analysis.

Sensitivity scenario	Purpose of scenario	Finding
AP2 2011 damage estimates	Test more recent estimates of damages per ton emissions from each location	Higher estimated damages, but key findings are robust. Increased potential for PEVs to lower damages versus HEVs in future grid
EASIUR damage estimates	Test an alternative method for estimating damages	Higher estimated damages, but key findings are robust. PEVs have comparable damages versus HEVs in future grid
High efficiency long-range BEV	Test the life cycle damages of a BEV with comparable performance to the other vehicles	Lower estimated damages but key findings are robust. BEVs have higher damages than CVs, HEVs, and PHEVs in the recent grid
Future with expensive coal, cheap natural gas	Test effect of possible future fuel prices	PEVs slightly increase damages versus HEVs when uncontrolled, reduce damages when controlled
Future with cheap coal, expensive natural gas	Test effect of possible future fuel prices	PEVs slightly reduce damages versus HEVs whether uncontrolled or controlled
Additional spinning reserves required for high wind case	Test effect of higher reserve requirements for intermittent wind	Key findings are robust
No transmission constraints	Test effect of transmission constraints and location of loads and responding generators	Key findings are robust
PEVs in urban areas only	Test effect of location of vehicles	Key findings are robust
SO_2 cap	Test effect of a binding SO_2 cap	SO_2 damages likely reduced, but not necessarily to zero since PEV charging may affect location of SO_2 emissions. PEVs likely reduce damages
CAFE leakage effect	Test effect of vehicle fleet-wide emissions enabled by PEV sales due to CAFE incentives for AFV adoption	PEVs increase overall US damages versus HEVs even in future grid and even with high wind penetration

and detailed in the supporting information. These include alternative assumptions for (1) emissions damage models, (2) BEV efficiency, (3) power grid characteristics that influence dispatch decisions, (4) PEV adoption patterns, and (5) policy effects on consequential life cycle implications. Our key findings are generally robust, though future fuel prices could affect the consequential emissions benefits of PEVs versus HEVs; binding SO_2 caps would likely make PEVs more competitive; and assessing a larger consequential life cycle scope that accounts for the effect of PEV adoption on each manufacturer's vehicle fleet emissions due to US corporate average fuel economy and GHG emission policy results in the conclusion that PEV adoption increases damages in all scenarios modeled (the policy extends through 2025) [29].

4. Discussion and conclusions

Using a consequential LCA approach for the vehicles and scenarios modeled, we find that (1) PEVs cause more damage than HEVs in the recent PJM grid, (2) PEVs cause comparable or slightly lower emissions than HEVs in the future PJM grid scenarios, and (3) utility-controlled (mostly nighttime) PEV charging tends to increase life cycle emissions compared to uncontrolled charging. However, (1) changes in future fuel prices could affect whether PEVs have higher or lower damages than HEVs in a future grid and whether controlled charging increases or decreases emissions, and (2) when the effect of US PEV adoption on automaker fleet emissions is accounted for, PEV adoption could increase emissions in all scenarios due to leakage effects in federal fuel economy standards [29, 34–36]. The SI includes a discussion about the limitations of the model beyond the scope of the sensitivity analysis.

Our results for the recent PJM grid are consistent with those from Tessum *et al* using the 2007 electricity mix [2]: PEVs have higher damages than gasoline vehicles in the recent grid. Michalek *et al* [1] found that PEVs with larger batteries cause more damage, which we also observe in our recent grid scenarios. None of the PEVs in our study, regardless of battery size, have lower damages than the HEV in the recent grid due to the large amounts of coal on the margin in PJM compared to the average mix used in Michalek *et al* [1]. The future of the grid past 2018, which is relevant for future PEV adoption, is expected to be lower-emitting than the recent grid, but consequential emissions from a grid far into the future are difficult to meaningfully project. Both Michalek *et al* [1] and Tessum *et al* [2] find that PEVs can reduce damages if charged with zero-emission electricity. This is a useful bounding case but not a scenario likely to occur soon, since even if wind, solar, nuclear, and hydroelectric power make up a much larger portion of the grid mix in the future, the consequential effect of PEV charging (the difference between grid operations when PEVs are present versus absent) is still primarily to increase generation from fossil fuel plants. Since most wind, solar, nuclear, and hydroelectric power is fully used in the absence of PEV load, PEV adoption will not cause an increase in generation from these plants. Only when low-emission plants would have been curtailed in the absence of PEVs can PEV adoption result in increased use of these plants. Thus, the consequential emissions of PEV charging are affected more by the mix of coal and natural gas plants in a region than by the amount of renewable or low-emission generation capacity.

The difference between coal and natural gas generation is significant. We show that even in one of the power systems in the country with the highest coal generation, PEVs could reduce transportation health and environmental damages in the near future, long before a zero-carbon electricity mix is achieved, due primarily to substitution of natural gas for coal on the margin.

While PEVs can double or triple air emission damages in the recent grid relative to HEVs, they could reduce damages in a future grid. However, we estimate that near future (~2018) potential air emissions benefits from PEV adoption in PJM are small relative to HEVs (or even negative when considering the net effect on the automaker's fleet under federal fuel economy policy). Nevertheless, electrification may offer a promising long term option to significantly reduce air emissions from the transportation sector compared to some other alternative transportation fuels, including biofuels and natural gas, that have been shown to offer small-to-no reductions in GHG emissions and could have unintended consequences like higher global food prices [30, 31]. Indeed, the logistics of regulating emissions from individual vehicles over their functional lives are more difficult than regulation of power plant emissions [37].

Continued regulation of the electricity system can increase the benefits of vehicle electrification, and consequential air emissions implications of PEV charging are already lower in many regions than in PJM [6]. While near-term benefits of PEV adoption in PJM are estimated to be small or negative, a transition of the transportation system could lead to long-term benefits outside the scope of this analysis, including greater benefits in other regions and future emissions savings enabled by a transition to electric vehicles as the electricity grid becomes cleaner and as public policy adjusts [32, 33].

Acknowledgments

We would like to thank to Nick Muller for his support with the APEEP and AP2 models as well as Peter Adams and Jinhyok Heo for their help with the EASIUR model. The Doris Duke Charitable Foundation, the Richard King Mellon Foundation, the Electric Power Research Institute, and the Heinz Endowment provided support for this research through the RenewElec project (www.renewelec.org). The National Science Foundation also provided financial support through a Graduate Research Fellowship, and Toyota Motor Corporation provided support through a financial gift. Findings and conclusions are the sole responsibility of the authors and may not represent the views of the funding entities.

References

[1] Michalek J *et al* 2011 Valuation of plug-in vehicle life-cycle air emissions and oil displacement benefits *Proc. Natl Acad. Sci. USA* **108** 16554–8

[2] Tessum C, Hill J and Marshall J 2014 Life cycle air quality impacts of conventional and alternative light-duty transportation in the United States *Proc. Natl Acad. Sci. USA* **111** 18490–5

[3] Eckelman M 2013 Life cycle assessment in support of sustainable transportation *Environ. Res. Lett.* **8** 021004

[4] Samaras C and Meisterling K 2008 Life cycle assessment of greenhouse gas emissions from plug-in hybrid vehicles: implications for policy *Environ. Sci. Technol.* **42** 3170–6

[5] Hawkins T, Singh B, Majeau-Bettez G and Strømman A H 2013 Comparative environmental life cycle assessment of conventional and electric vehicles *J. Ind. Ecology* **17** 53–64

[6] Tamayao M, Michalek J J, Hendrickson C and Azevedo I 2015 Regional variability and uncertainty of electric vehicle life cycle CO_2 emissions across the United States *Environ. Sci. Technol.* **49** 8844–55

[7] Silas-Evans K, Azevedo I and Morgan M G 2012 Marginal emissions factors for the US electricity system *Environ. Sci. Technol.* **46** 4742–8

[8] Graff Zivin J, Kotchen M and Mansur E 2014 Spatial and temporal heterogeneity of marginal emissions: implications for electric cars and other electricity-shifting policies *J. Econ. Behav. Organ.* **107** 248–68
[9] Sioshansi R and Denholm P 2010 The value of plug-in hybrid electric vehicles as grid resources *Energy J.* **31** 1–10
[10] Peterson S B, Whitacre J F and Apt J 2011 Net air emissions from electric vehicles: the effect of carbon price and charging strategies *Environ. Sci. Technol.* **45** 1792–7
[11] Choi D G, Kreikebaum F, Thomas V and Divan D 2013 Coordinated EV adoption: double-digit reductions in emissions and fuel use for $40/vehicle-year *Environ. Sci. Technol.* **47** 10703−10707
[12] Weis A, Michalek J, Jaramillo P and Lueken R 2015 Emissions and Cost implications of controlled electric vehicle charging in the US PJM interconnection *Environ. Sci. Technol.* **49** 5813–9
[13] Ma H, Balthasar F, Tait N, Riera-Palou X and Harrison A 2012 A new comparison between the life cycle greenhouse gas emissions of battery electric vehicles and internal combustion vehicles *Energy Policy* **44** 160–73
[14] Thomas S 2012 US marginal electricity grid mixes and EV greenhouse gas emissions *Int. J. Hydrog. Energy* **37** 19231–40
[15] Onat N C, Kucukvar M and Tatari O 2015 Conventional, hybrid, plug-in hybrid or electric vehicles? State-based comparative carbon and energy footprint analysis in the United States *Appl. Energy.* **150** 36–49
[16] Hadley W S and Tsvetkova A 2008 *Potential Impacts of Plug-in Hybrid Electric Vehicles on Regional Power Generation* Oak Ridge, Tennessee (http://web.ornl.gov/info/ornlreview/v41_1_08/regional_phev_analysis.pdf)
[17] FERC (http://ferc.gov/market-oversight/mkt-electric/pjm.asp)
[18] Lueken R and Apt J 2014 The effects of bulk electricity storage on the PJM market *Energy Syst.* **5** 677–704
[19] EPA 2010 National Electric Energy Data System (NEEDS) v.4.10 (www.epa.gov/airmarkets/integrated-planning-model-ipm-base-case-v410)
[20] Environmental Protection Agency 2018 Parsed Results Detailed Output Files (http://epa.gov/airmarkets/progsregs/epa-ipm/BaseCasev513.html)
[21] NREL 'Eastern Wind Dataset' (http://nrel.gov/electricity/transmission/eastern_wind_dataset.html) (accessed 7 July 2011)
[22] Argonne National Laboratory 2013 GREET 1 (https://greet.es.anl.gov/)
[23] Argonne National Laboratory 2013 GREET 2 (https://greet.es.anl.gov/)
[24] EPA 2012 Tesla Model S,' fueleconomy.gov
[25] US Department of Transportation, Federal Highway Administration 2009 National Household Travel Survey 2009 (http://nhts.ornl.gov) (accessed 6 October 2011)
[26] Weis A, Jaramillo P and Michalek J 2014 Estimating the potential of controlled plug-in hybrid electric vehicle charging to reduce operational and capacity expansion costs for electric power systems with high wind penetration *Appl. Energy* **115** 190–204
[27] Interagency Working Group on the Social Cost of Carbon Technical Support Document:—Technical Updated of the Social Cost of Carbon for Regulatory Impact Analysis—Under Executive Order 12866 (http://whitehouse.gov/sites/default/files/omb/assets/inforeg/technical-update-social-cost-of-carbon-for-regulator-impact-analysis.pdf)
[28] Muller N AP2 (APEEP) Model (https://sites.google.com/site/nickmullershomepage/home/ap2-apeep-model-2)
[29] Jenn A, Michalek J J and Azevedo I L Unintended consequences: Why US alternative fuel vehicle adoption increases gasoline consumption and greenhouse gas emissions (in review)
[30] Mullins K A, Griffin W M and Matthews H S 2011 Policy Implications of uncertainty in modeled life-cycle greenhouse gas emissions of biofuels *Environ. Sci. Technol.* **45** 132–8
[31] Tong F, Jaramillo P and Azevedo I M L 2015 Comparison of life cycle greenhouse gases from natural gas pathways for light-duty vehicles *Energy Fuels* **29** 6008–18
[32] Greene D, Park S and Liu C 2013 *Analyzing the Transition to Electric Drive in California* The Howard H. Baker Jr Center for Public Policy (www.theicct.org/sites/default/files/publications/Transition-to-Electric-Drive-2013-report.FINAL_.pdf)
[33] National Research Council of the National Academies 2013 *Transitions to Alternative Vehicles and Fuels* (www.nap.edu/catalog/18264/transitions-to-alternative-vehicles-and-fuels)
[34] Gecan R, Kile J and Beider P 2012 *Effects of Federal Tax Credits for the Purchase of Electric Vehicles* US Congressional Budget Office pp 1–36 (www.cbo.gov/sites/default/files/cbofiles/attachments/09-20-12-ElectricVehicles_0.pdf)
[35] Goulder L, Jacobsen M and Van Benthem A 2012 Unintended consequences from nested state and federal regulations: the case of the Pavley greenhouse-gas-per-mile limits *J. Environ. Econ. Manage.* **63** 187–207
[36] Jenn A 2014 Advanced and alternative fuel vehicle policies: regulations and incentives in the United States *PhD Dissertation* Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA
[37] Beaton S P *et al* 1995 On-road vehicle emissions—regulations, costs, and benefits *Science* **268** 991–3

Environmental Research Letters

OPEN ACCESS

RECEIVED
13 December 2015

REVISED
2 March 2016

ACCEPTED FOR PUBLICATION
16 March 2016

PUBLISHED
7 April 2016

LETTER

Effect of regional grid mix, driving patterns and climate on the comparative carbon footprint of gasoline and plug-in electric vehicles in the United States

Tugce Yuksel[1], **Mili-Ann M Tamayao**[2,4], **Chris Hendrickson**[2,3], **Inês M L Azevedo**[2] **and Jeremy J Michalek**[1,2,5]

1 Department of Mechanical Engineering, Carnegie Mellon University, 324 Scaife Hall, Pittsburgh, PA 15213, USA
2 Department of Engineering and Public Policy, Carnegie Mellon University, 129 Baker Hall, Pittsburgh, PA 15213, USA
3 Department of Civil and Environmental Engineering, Carnegie Mellon University, 119 Porter Hall, Pittsburgh, PA 15213, USA
4 Present affiliation: Department of Industrial Engineering and Operations Research, University of the Philippines-Diliman, Quezon City1101, Philippines.
5 Author to whom any correspondence should be addressed.

E-mail: jmichalek@cmu.edu

Keywords: electric vehicle, life cycle assessment, greenhouse gas emissions, carbon footprint

Supplementary material for this article is available online

Abstract

We compare life cycle greenhouse gas (GHG) emissions from several light-duty passenger gasoline and plug-in electric vehicles (PEVs) across US counties by accounting for regional differences due to marginal grid mix, ambient temperature, patterns of vehicle miles traveled (VMT), and driving conditions (city versus highway). We find that PEVs can have larger or smaller carbon footprints than gasoline vehicles, depending on these regional factors and the specific vehicle models being compared. The Nissan Leaf battery electric vehicle has a smaller carbon footprint than the most efficient gasoline vehicle (the Toyota Prius) in the urban counties of California, Texas and Florida, whereas the Prius has a smaller carbon footprint in the Midwest and the South. The Leaf is lower emitting than the Mazda 3 conventional gasoline vehicle in most urban counties, but the Mazda 3 is lower emitting in rural Midwest counties. The Chevrolet Volt plug-in hybrid electric vehicle has a larger carbon footprint than the Prius throughout the continental US, though the Volt has a smaller carbon footprint than the Mazda 3 in many urban counties. Regional grid mix, temperature, driving conditions, and vehicle model all have substantial implications for identifying which technology has the lowest carbon footprint, whereas regional patterns of VMT have a much smaller effect. Given the variation in relative GHG implications, it is unlikely that blunt policy instruments that favor specific technology categories can ensure emission reductions universally.

1. Introduction

Past studies have shown that life cycle plug-in electric vehicle (PEV) emissions depend heavily on the assumed electricity grid mix [1–7], driving patterns (including drive cycle and distance) [8–10] and climate (including ambient temperature) [7, 11]. These factors vary regionally, so PEV emissions implications also vary regionally. Several studies have assessed regional differences in PEV emissions incorporating subsets of these factors [2, 4–7, 11–15]—with most focused on regional grid mix, but no study has accounted for the combined influence of consequential grid emissions, driving patterns, and temperature heterogeneity in assessing regionally-specific life cycle implications of PEVs in the US. In table 1 we summarize studies that make regional comparisons of PEV emissions in the United States. Key factors that differentiate these studies include:

1.1. Life cycle scope

Existing studies assessing PEV emissions have different life cycle scopes, which may include or exclude each of the following: vehicle and battery manufacturing emissions; gasoline extraction, processing, transportation, and fuel combustion emissions; power

Table 1. Summary of published studies assessing the regional variation in electrified vehicle GHG emissions in the United States.

Study	Vehicle types	Regional resolution	Life cycle scope	Electricity source and emissions	Utility factor or VMT pattern	Driving conditions	Temperature
EPRI-NRDC (2007) [14]	PHEV	NERC regions	**Use Phase** Electricity upstream and generation; gasoline upstream and combustion	**Consequential** Bottom-up modeled emissions (573 g $CO_2e\ kWh^{-1}$ in 2010; 97–412 g $CO_2e\ kWh^{-1}$ in 2050)	**Homogeneous** ($PHEV_{10}$: 0.12 $PHEV_{20}$: 0.49 $PHEV_{40}$: 0.66)	**Homogeneous** Federal Urban Driving Schedule (FUDS)	**Ignored**
Hadley and Tsvetkova (2009) [15]	PHEV	13 NERC subregions	**Partial Use Phase** Electricity generation; gasoline combustion	**Consequential** Bottom-up approach using ORCED model assuming 25% PHEV market penetration by 2020	**Homogeneous** Not clear	**Homogeneous** Three load levels assumed per vehicle—1.5 kW, 2 kW, and 6 kW	**Ignored**
Anair and Mahmassani (2012) [4]	ICV, HEV, PHEV, BEV	eGRID subregion	**Use Phase** Electricity upstream and generation; gasoline upstream and combustion	**Attributional** Average regional generation covering transmission and upstream loss (286–983 gCO2e kWh^{-1})	**Homogeneous** (Chevrolet Volt: 0.64)	**Homogeneous** EPA combined driving cycle[a]	**Ignored**
MacPherson *et al* (2012) [17]	PHEV	NERC regions, NERC subregions and states	**Life Cycle**	**Attributional** Average regional and state based emissions from EPA eGRID2010 database	**Homogenous:** $PHEV_{35}$: 0.635, and **Regional:** NERC region based utility factors estimated based on NHTS.	**Homogeneous** EPA combined driving cycle	**Ignored**
Thomas (2012) [19]	HEV, PHEV, BEV	13 NERC subregions	**Use Phase** Electricity upstream and generation from GREET; gasoline well-to-wheel using GREET (2001)	**Consequential** Average marginal emissions from Hadley and Tsvetkova (2009)	**Not clear**	**Not clear**	**Ignored**
Yawitz *et al* (2013) [13]	HEV, PHEV, BEV	State	**Life Cycle**	**Attributional** Average state generation	**Homogeneous** PHEV: 0.5	**Homogeneous** EPA 2013	**Ignored**
Graff Zivin *et al* (2014) [5]	ICV, HEV, PHEV, BEV	eGRID subregion	**Partial Use Phase** Electricity generation; gasoline combustion	**Consequential** Marginal NERC emissions considering interregional trading	**Homogeneous** 35 mi day^{-1}	**Homogeneous** EPA combined city/highway	**Ignored**

Table 1. (Continued.)

Study	Vehicle types	Regional resolution	Life cycle scope	Electricity source and emissions	Utility factor or VMT pattern	Driving conditions	Temperature
Onat *et al* (2015) [6]	ICV, HEV, PHEV, BEV	13 NERC subregions	**Life Cycle**	**Consequential** Marginal emissions from Thomas (2012) [19] which is based on ORCED model	**Regional** State based utility factors	**Homogeneous** EPA combined	**Ignored**
Tamayao *et al* (2015) [2]	ICV, HEV, PHEV, BEV	NERC region	**Life Cycle**	**Consequential** Compares Graff Zivin *et al* (2014) and Siler-Evans *et al* (2012) marginal emission factors by NERC region and average state, eGRID subregion, and NERC emission factors	**Homogeneous** US NHTS (2009) national distribution	**Homogeneous** EPA combined	**Ignored**
Yuksel and Michalek (2015) [11]	BEV	NERC region	**Partial Use Phase** Electricity generation; gasoline combustion	**Consequential** Compares Graff Zivin *et al* (2014) and Siler-Evans *et al* (2012) marginal emission factors by NERC region.	**Homogeneous** US NHTS (2009) national distribution	**Homogeneous** Efficiency based on FleetCarma on-road data [28]	**Regional** Based on FleetCarma data for Nissan Leaf and regional temperature data
Nealer *et al* (2015) [18]	BEV	eGRID subregions	**Life Cycle**	**Attributional** Average emission rate for generators located in each subregion.	**Homogeneous**	**Homogeneous** EPA combined city/highway	**Ignored**
Archsmith *et al* (2015) [20]	ICV, BEV	NERC regions	**Life Cycle**	**Consequential** Regression-based marginal emission estimates for current, average emission rates for future	**Regional** Based on regional NHTS data	**Homogeneous** Based on GREET	**Regional** Based on data from [26–27]
This study	ICV, HEV, PHEV, BEV	County-level estimates based on highest-resolution data available for each factor	**Life Cycle**	**Consequential** Compares Graff Zivin *et al* (2014) and Siler-Evans *et al* (2012) marginal emission factors by NERC region.	**Regional** NHTS (2009) state distribution	**Regional** EPA city, highway or combined based on county urbanization level	**Regional** Based on ANL temperature-controlled laboratory test data and regional temperature data.

[a] The US Department of Energy define three driving conditions: city—'urban driving, in which a vehicle is started in the morning (after being parked all night) and driven in stop-and-go traffic'; highway—'a mixture of rural and interstate highway driving in a warmed-up vehicle, typical of longer trips in free-flowing traffic'; (3) combined—'combination of city driving (55%) and highway driving (45%)' [48].

plant emissions from electricity generation for vehicle charging; power plant fuel feedstock extraction, production and transportation emissions; and end of life emissions. Several of the studies shown in table 1 only include emissions related to vehicle use or a subset of the emissions related to vehicle use (e.g., vehicle tailpipe emissions and power plant smokestack emissions), leading to incomplete assessments. Life cycle studies suggest that emissions implications from sources other than tailpipe and power plant emissions can comprise one fifth to one third of vehicle life cycle greenhouse gas (GHG) emissions [2, 6, 16, 17], so addressing the full life cycle can be important for comprehensive comparisons.

1.2. Electricity sources and emissions

Critical to assessing life cycle emissions of PEVs are the sources of energy used to generate electricity and their efficiencies [1, 2, 4, 5, 7, 12, 13]. While some studies use an attributional life cycle approach in which they assign to the PEV the average emission rates for power plants in the same state or power grid region where it is charged [4, 13, 17, 18], other studies take a consequential life cycle approach, estimating the change in grid emissions resulting from new PEV charging in a region [5, 7, 14, 15, 19]. The latter is appropriate for assessing the emissions implications of a policy intervention. One empirical approach to estimating consequential emissions of PEV charging is to estimate marginal emission factors using historical data. Several studies have conducted regressions on past data to estimate marginal emission rates for US grid regions [5, 20, 7], though Alexander *et al* (2015) warn that regional marginal emissions can be difficult to identify because of interregional trade [21]. Tamayao *et al* (2015) [2] show that differences between average and marginal emission factors can affect whether PEVs are estimated to be higher or lower emitting than efficient gasoline vehicle models. In some cases, the uncertainty is such that one is not able to conclude whether the emissions from PEVs are larger or smaller than efficient gasoline vehicle models.

1.3. Driving patterns

Driving conditions (specifically, driving cycle—the trajectory of vehicle velocity over time) can affect the relative vehicle efficiency of PEVs and conventional gasoline vehicles differently and thus substantially affect the relative economic and environmental benefits of electrified vehicles. For instance, PEVs can offer substantial economic and GHG benefits over conventional vehicles (CVs) for stop-and-go city driving while offering fewer environmental benefits at a higher cost premium for highway cruising [8]. Patterns of driving distance also matter, particularly for PHEVs, which use a mix of gasoline and electricity for propulsion. For example, longer driving distances lead to higher petroleum and total energy use [9, 10], and the shorter distances traveled by urban drivers result in higher PHEV utility factors [22]. As shown in table 1, most existing studies have modeled regional heterogeneity of electricity source but ignore regional differences in driving distance distributions and driving conditions that affect vehicle efficiency.

1.4. Temperature

Most studies ignore the regional effect of ambient temperature. However, temperature has an important effect on vehicle efficiency due to heating, ventilation, and air conditioning (HVAC) use and temperature-related battery efficiency effects. Indeed, compared to mild regions, Yuksel and Michalek (2015) [11] estimate that battery electric vehicles (BEVs) can consume an average of 15% more energy in hot and cold regions of the US. Similarly, Neubauer and Wood (2014) [23] estimate that HVAC use can increase energy consumption by 24% in cold climates, and Kambly and Bradley (2014) [24, 25] note that HVAC use can decrease BEV range depending on the region and time of day; and Meyer *et al* (2012) [26] observe a 60% drop in range in −20 °C lab tests with maximum climate control use. Archsmith *et al* (2015) [7] use vehicle test data from Meyer *et al* (2012) [26] and Lohse-Busch *et al* (2013) [27] to argue that temperature can have as large an effect on electric vehicle charging emissions as regional grid mix, and in a working paper Holland *et al* (2015) [12] adjust vehicle efficiency regionally to account for temperature effects in estimating air pollution damages.

To assess the combined effect of these regional factors, we develop and apply a model that integrates the effects of electricity source, driving patterns, and temperature with a comprehensive life cycle scope to characterize regional GHG emissions from electricity and gasoline light-duty vehicles.

2. Data and methods

We perform a comparative life cycle assessment of the CO_2 emissions across five existing vehicle models summarized in table 2. These vehicle models represent CVs, hybrid electric vehicles (HEVs), plug-in electric vehicle (PHEVs), and BEVs, and they were selected based on availability of Argonne National Laboratory vehicle test efficiency data at high, low, and moderate test chamber temperatures [29].

Figure 1 summarizes the framework used in this work. We start by assigning driving conditions to each county based on urbanization level; we assign vehicle miles traveled (VMT) patterns to counties based on data from the National Household Travel Survey (NHTS) for the corresponding state; and we assign marginal grid emission factors for each North American Electric Reliability Corporation (NERC) region to the counties that lie in that region. We then estimate the energy consumption rate for each vehicle based on

Table 2. Vehicle models considered.

Vehicle model	Type	Model year	Battery energy capacity	
			Nominal (kWh)	Usable (kWh)
Nissan Leaf	BEV	2013	24	21
Chevy Volt	PHEV (EREV)	2013	16.5	10.8
Toyota Prius PHEV	PHEV (blended)	2013	4.4	2.7
Toyota Prius	HEV	2010	—	—
Mazda 3 (with i-ELOOP)	CV	2014	—	—

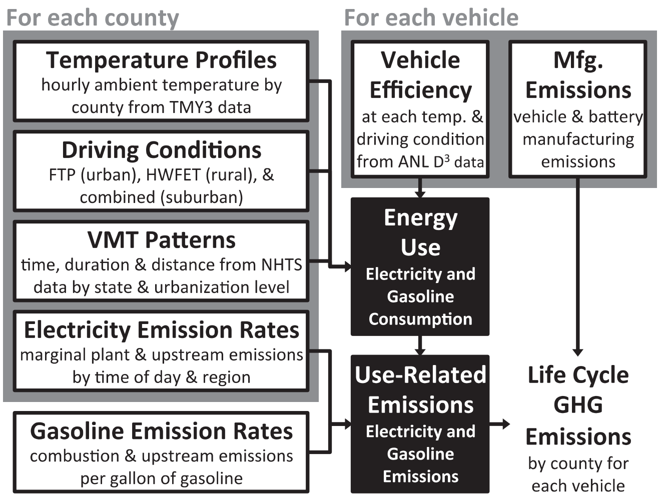


Figure 1. Framework for the analysis.

Argonne National Laboratory's Downloadable Dynamometer Database (D^3) temperature-controlled chamber vehicle test data together with information on temperature, drive cycle, and VMT patterns for each county. We use energy consumption and VMT patterns to compute timing and duration of vehicle charging. Finally, we estimate life cycle CO_2 emissions for each vehicle type and location by adding vehicle and battery manufacturing emissions, gasoline combustion and upstream emissions (based on computed gasoline consumption), and electricity production and upstream emissions (based on computed electricity consumption, timing, and location).

We use county-level data when such resolution exists, and we use regional data where we lack county-level resolution. We perform sensitivity analysis to test implications of several factors and assumptions and to test robustness of our results. We explain each of these modules in the following sections with additional detail provided in the SI.

2.1. Vehicle energy efficiency

For each vehicle model in table 2 we estimate how vehicle energy efficiency changes with driving cycle and temperature. We use the D^3 database from Argonne National Laboratory's Advanced Powertrain Research Facility [29], which provides dynamometer test data for several vehicle models. D^3 provides energy efficiency estimates at three different temperatures (20° F, 72° F and 95° F) and for three different standard test driving cycles (the urban dynamometer driving schedule (UDDS) cycle, the US06 cycle, and the highway fuel economy test (HWFET) cycle) [30]. During the tests at 20° F and 95° F, the climate control is set to keep the cabin temperature at 72° F. These '2-cycle' tests, used in federal regulatory compliance calculations, are known to produce optimistic fuel consumption results relative to on-road driving, resulting in lower than actual emission estimates [31]. We use linear interpolation between each measured point, and we avoid extrapolation below 20° F and above 95° F (instead holding the efficiency estimate fixed at the corresponding extremum for temperatures outside the measured ranges). The measured efficiency estimates account for charging losses.

2.2. VMT patterns

Daily trip length and timing for light-duty vehicles in each county is drawn from the distribution of trips in the NHTS [32] from all counties from the same state from a set of 76 149 total vehicles (we filter the dataset to private light-duty vehicles only and exclude the data points that are reported by members of the household other than the driver). Trip details are used to account for the ambient temperature effect (as temperature varies through the day) and to assess when the vehicle

Table 3. Assumptions and data sources used for each life cycle stage.

Emissions source	Estimate(s) used	Data source
Vehicle manufacturing (including battery)	18 g mi^{-1} CV	GREET (2013) [35] and Tamayao *et al* (2015) [2]
	16 g mi^{-1} HEV	
	41 g mi^{-1} PHEV-EREV	
	22 g mi^{-1} PHEV-blended	
	51 g mi^{-1} BEV	
Gasoline combustion	8655 gCO_2 gal^{-1} gasoline	Average of values from EPA (2014) [36] and Venkatesh *et al* (2011) [37]
Gasoline production and transportation	2400 gCO_2 gal^{-1} gasoline	Average of values from Venkatesh *et al* (2011) [37] and GREET (2013) [35]
Electricity generation	430–932 $kgCO_2eq$ MWh^{-1}	Siler-Evans *et al* (2012) [20]
Electricity upstream	38–107 kg CO_2 MWh^{-1}	Tamayao *et al* (2015) [2] (estimated based on Siler-Evans *et al* (2012) [25] Graff Zivin *et al* (2014) [5], Venkatesh *et al* (2011) [37], Venkatesh *et al* (2011) [38], and US EPA (2009) [39])

is available for charging. We test alternative assumptions in the sensitivity analysis.

2.3. Driving conditions

For urban counties we use the UDDS test results; for rural counties we use the HWFET cycle results; and for outlying (suburban) counties we use the combined results to represent the dominant driving conditions in each case. We test alternative assumptions in the sensitivity analysis.

2.4. Charging profile

We assume convenience charging, i.e., charging starts as the last trip of the day ends. We estimate the charging duration based on the daily energy consumption of each vehicle. We test alternative assumptions in the sensitivity analysis.

2.5. Temperature

We use the Typical Meteorological Year (TMY3) Database from the National Renewable Energy Laboratory [33] which provides hourly ambient temperature data for a typical meteorological year for 1011 locations in the continental United States. We use a triangulation-based linear spatial interpolation method [34] to estimate temperature profiles at the center of each county. In the sensitivity analysis, we assess the effect of ignoring temperature on our results.

2.6. Emission factors

For electricity emissions associated with PEV charging, we use the 2011 marginal emission factors from Siler-Evans *et al* [20], which are based on regressions of empirical, historical changes in power plant emissions with respect to changes in generation within each NERC region. We examine this choice in more detail in the discussion section and test alternative assumptions in the sensitivity analysis.

Table 3 summarizes emissions estimates associated with manufacturing and assembly of vehicles and lithium-ion battery packs; gasoline production, transport and combustion; and electricity upstream, production, transmission, and distribution.

3. Results and discussion

Figure 2 summarizes the increase or decrease in life cycle GHG emissions from driving a 2013 Nissan Leaf BEV, a 2013 Chevrolet Volt PHEV, or a 2013 Toyota Prius PHEV relative to the most efficient gasoline vehicle in the market—the Toyota Prius HEV (modeled here using data from a 2010 HEV Prius)—and relative to a CV of comparable size—the Mazda 3. A map of county urbanization level is provided in the SI, since urbanization level determines drive cycle.

The Nissan Leaf BEV produces lower life cycle GHG emissions than the Prius HEV in urban counties of Texas, Florida, and much of the southwestern US. In most of the rest of the country the Leaf increases GHG emissions relative to the Prius HEV, with those increases being most notable in the Midwest and in the South. This is due to the combined effect of grid carbon intensity, highway driving, and regional temperature. In particular, the Northern Midwest has a combination of a coal-heavy electricity grid, rural counties (with an assumed highway driving cycle), and cold weather that all contribute to higher relative emissions for the BEV.

The Chevrolet Volt PHEV has higher life cycle emissions than the Prius HEV in all counties. This is because the Volt consumes more gasoline per mile in charge-sustaining mode (after the battery is depleted) than the Prius HEV, and it consumes more electricity per mile than the Leaf in charge-depleting (CD) mode (when the battery is charged) at high temperatures. Further, in cold weather the Volt consumes both gasoline and electricity in CD mode. Comparison of electricity and gasoline consumption for different vehicles is provided in the SI (section 4).

The PHEV Prius produces lower life cycle GHG emissions than the HEV Prius in Texas, Florida, and

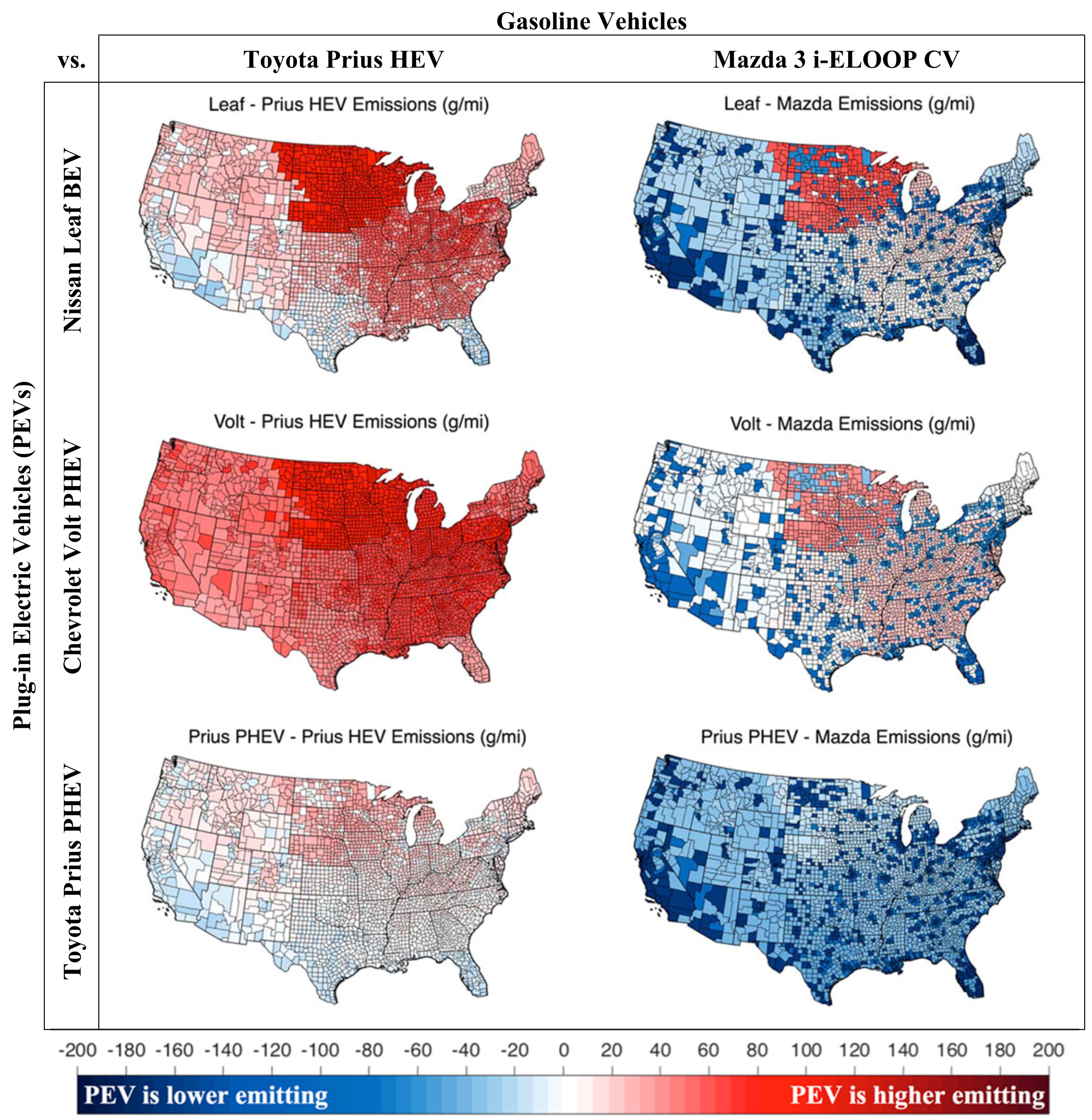


Figure 2. Estimated difference in life cycle GHG emissions ($gCO_2eq\ mi^{-1}$) of selected plug-in electric vehicles (2013 Nissan Leaf BEV, 2013 Chevrolet Volt PHEV, and 2013 Prius PHEV) relative to selected gasoline vehicles (2010 Prius HEV and 2014 Mazda 3). In each case blue indicates that the PEV has lower GHG emissions than the gasoline vehicle and red indicates that the PEV has higher GHG emissions than the gasoline vehicle.

the southwestern US as well as in most urban areas, but it produces higher emissions in many rural areas across the country—especially in the Northern Midwest. This is because the PHEV Prius consumes less gasoline than the HEV Prius in city driving conditions and more gasoline than the HEV Prius in highway driving conditions. Differences between the HEV Prius and the PHEV Prius are generally less pronounced than those comparing the HEV Prius to the Volt or the Leaf.

In the right-hand column in figure 2 we provide a similar analysis using a conventional gasoline vehicle, the 2014 Mazda 3 (with i-ELOOP), with EPA-rated combined (5-cycle) fuel efficiency of 32 mpg as the reference vehicle in place of the HEV Prius. The i-ELOOP is an energy recovery braking system intended to capture a portion of the benefits that HEVs and PEVs capture in regenerative braking to displace accessory load without a full hybrid system. Relative to the Mazda 3, we find that (1) the Leaf reduces GHG emissions in urban counties across the US as well as suburban and rural counties in Texas, Florida, the Western US, and New England while increasing GHG emissions in the rural Midwest; (2) the Volt reduces GHG emissions in urban counties across the US while increasing GHG emissions in rural counties of the Midwest and the South; and (3) the Prius PHEV reduces emissions in all counties. In all three cases the GHG emission reductions in urban counties can be substantial.

Figure 3 shows the breakdown of life cycle CO_2 emissions for each vehicle in various selected counties

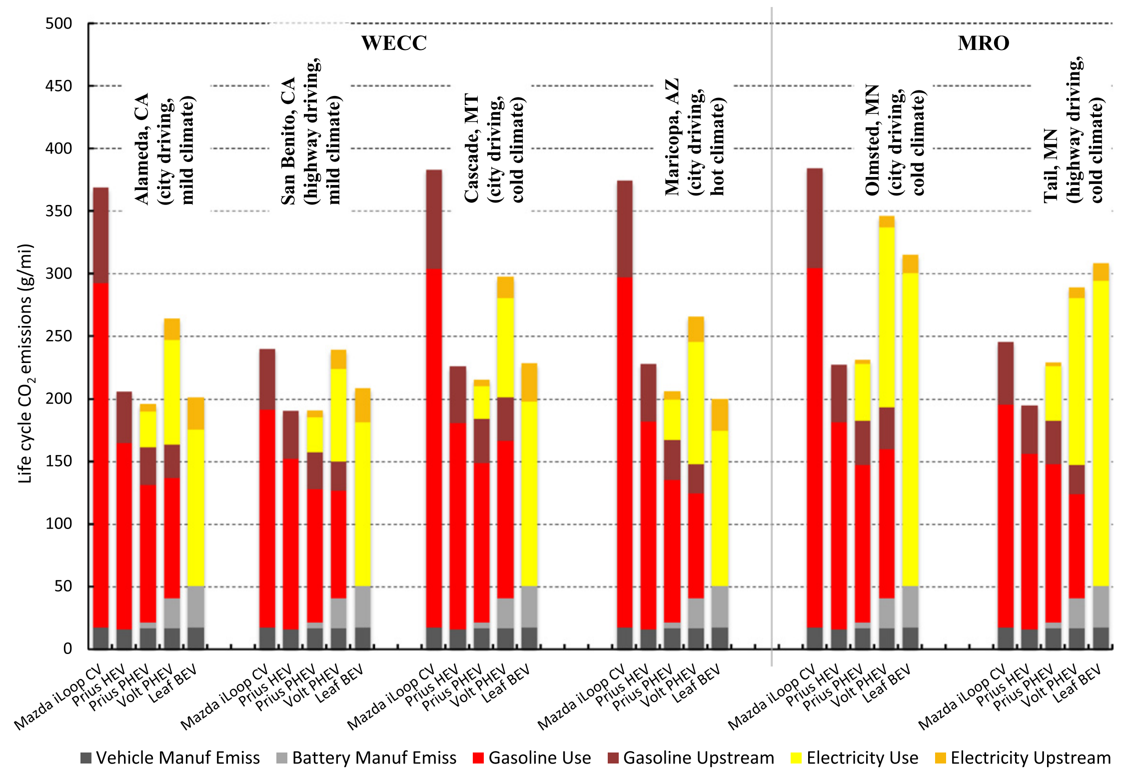


Figure 3. Life cycle CO_2 emissions in $gCO_2eq\ mi^{-1}$ in selected counties. Vehicles are ordered from lowest to highest degree of electrification.

from two NERC regions: the Western Electricity Coordinating Council (WECC) and the Midwest Reliability Organization (MRO), which have, respectively, the lowest and highest electricity generation CO_2 emissions factors in the continental US. The counties selected within those regions also have diverse climate and urbanization levels. Tailpipe and power plant emissions make up 64%–80% of life cycle GHG emissions in these examples. Batteries are less efficient when cold, and so are engines, but gasoline vehicles are able to use waste heat from the engine to heat the cabin, while BEVs and EREV PHEVs need to draw energy from the battery to heat the cabin, so PEVs tend to have larger energy penalties in cold weather regions than conventional gasoline vehicles.

The following conclusions can be made from figure 3:

- The effects of regional climate and grid mix on emissions become more important for vehicles with higher degrees of electrification. We find all vehicles have higher emissions in Minnesota, a colder state, compared to California. However, the increase in emissions is largest for the Leaf BEV, whereas only a slight increase is observed with Mazda 3 CV.
- In contrast, the effect of driving cycle on emissions becomes more prominent for vehicles with lower degrees of electrification. In counties with similar climate conditions and grid mix, we observe that the biggest change in emissions with highway driving compared to city driving occurs with Mazda 3.
- Hot temperatures in Arizona do not increase the emissions from the Leaf significantly relative to mild climate counties in California—an apparent contradiction to Yuksel and Michalek [11], who show a 22% increase in Leaf emissions in hot regions of Arizona compared to coastal California. The primary reason is that the laboratory data used in this study suggest lower energy consumption at high temperatures compared to real world data used in Yuksel and Michalek [11]. Further discussion of this issue is provided in the SI, section 4.

4. Sensitivity analysis

Details regarding the sensitivity analysis can be found in the SI, and table 4 summarizes key findings. Overall, we find that ignoring regional heterogeneity of temperature or driving conditions (city/highway) affects carbon footprint technology comparisons substantially in some regions, whereas urban/rural heterogeneity of VMT patterns has a negligible effect. We also find, consistent with prior work [2], that delayed charging increases the GHG emissions associated with PEVs in most regions and reduces the potential for emissions savings when compared to gasoline vehicles.

Table 4. Summary of findings from the sensitivity analysis.

Sensitivity case	Change from base case	Purpose	Finding
Homogeneous temperature	Vehicle efficiency at 72° F used for all counties all year	Test importance of temperature effect	Temperature effect substantially changes comparison results for northern states
Homogeneous driving conditions	Vehicle efficiency on combined UDDS/HWFET used for all counties	Test importance of drive cycle	Drive cycle affects the relative benefits of PEVs versus HEVs (and especially versus CVs). Without differentiated drive cycles, urban counties are not distinct from nearby rural counties.
VMT clustered by state and urbanization level	Each county's VMT distribution is drawn from all NHTS data from the same state and urbanization level	Test importance of differences in urban/rural driving distance	Using MSA level VMT does not change the results significantly. The maximum change is around 2 g mi^{-1}.
Delayed charging	Each PEV's charging schedule begins at midnight, rather than upon arrival at home	Test importance of charge timing	Delayed charging increases GHG emissions of PEVs in most of the country and reduces competitiveness with the HEV.

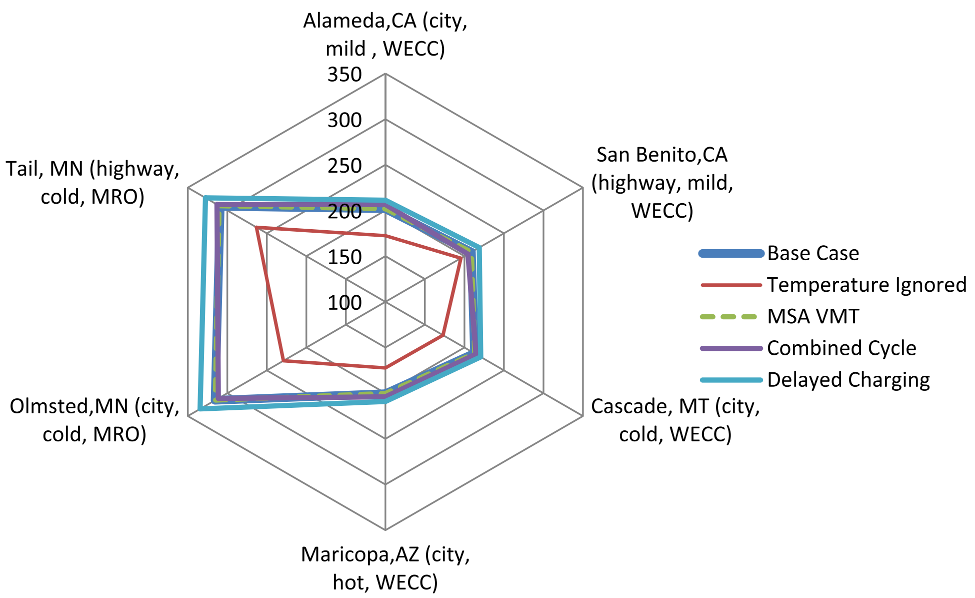


Figure 4. Radar chart showing Nissan Leaf life cycle emissions in $gCO_2eq\ mi^{-1}$ from different cases in selected counties.

Figure 4 summarizes life cycle GHG emission results for the Nissan Leaf in six counties. The Minnesota counties, which have both cold weather and the most carbon-intensive electricity grid region, have notably higher life cycle emissions than other counties, and the sensitivity case ignoring temperature has the largest effect on results.

Figure 5 summarizes the maximum change in GHG emissions per mile for a Nissan Leaf across all counties for each NERC region between the base case scenario and each sensitivity scenario. Ignoring temperature has the largest effect, reducing emissions estimates by up to 97 $gCO_2eq\ mi^{-1}$, while ignoring differences in drive cycle can increase emissions in some counties by up to 8 $gCO_2eq\ mi^{-1}$ (drive cycle affects CV efficiency more than PEV efficiency). Delayed charging can increase Leaf emissions by up to 21 $gCO_2eq\ mi^{-1}$, while use of MSA-level VMT patterns changes results less than 3 $gCO_2eq\ mi^{-1}$.

5. Limitations

Where possible, our analysis uses the most recent data available at the highest resolution available to account consistently for regional effects of grid emissions, driving patterns, and temperature on life cycle GHG emissions of PEVs and gasoline vehicles. However, there are several limitations regarding the data that should be understood when interpreting our results:

5.1. Regional grid emissions

The marginal emissions estimates used in this analysis are based on regressions for year 2011 and may not capture changes that may occur in the grid due to changes in policies, fuel prices, economic conditions or other factors. It is generally expected that GHG grid emission rates will decline over time, during the period that PEVs are being adopted and used. However, consequential (marginal) emissions from new load do

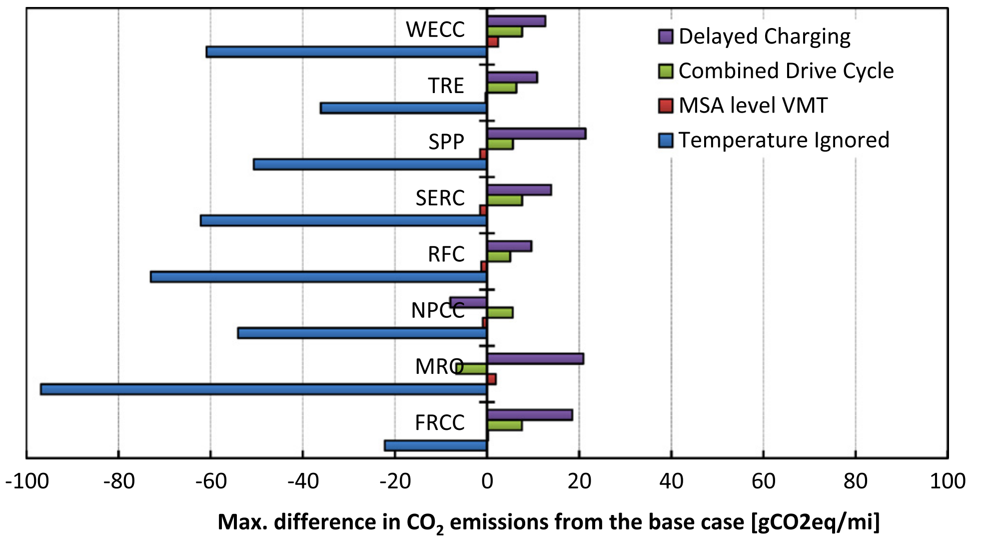


Figure 5. Maximum change in emissions for a Nissan Leaf relative to the base case. The maximum difference is observed in a different county for each case.

not decrease linearly with average grid emission rates. Because marginal emissions come primarily from fossil fuel plants, the mix of natural gas versus coal on the margin primarily determines the consequential emissions of new PEV charging. If the regions that are currently relying on coal at the margin start switching to natural gas generation used at the margin, then the amount of carbon dioxide savings from vehicle electrification will increase, and we may expect the more emissions-intensive areas of the country to look more like the less emissions-intensive areas in the future. Also, while we discuss county-level differences, we implicitly assume that within each NERC region all counties have identical marginal emission factors. Since the electric grid is heavily interconnected, it is difficult to attribute emissions to load changes at county-level resolution. In practice, it may be the case that adding PEV load in some areas of a NERC region could have different emission implications than adding the same load in a different area of the same NERC region.

5.2. Driving patterns

Our summary maps assign the UDDS test results to urban counties and the HWFET test results to rural counties, but in practice driving conditions are heterogeneous in all counties. Also, importantly, on-road driving conditions differ substantively from these two laboratory tests, which are known to produce optimistic fuel efficiency estimates due to their relatively mild drive cycle demands. Driving distances also may vary for different counties in a state, but we lump counties together when estimating driving distance distributions because we lack data resolution to identify driving distance distributions for individual counties. The NHTS data set provides information on the trips taken by each surveyed US vehicle on a single survey day and does not include day-to-day variability for each vehicle. In this study, we average over the vehicle profiles to assess implications for average driving distances and we assume these daily profiles are identical over the year. In practice the driving profiles of PEV adopters may differ from the general population.

5.3. Temperature

We treat temperature as the only factor affecting vehicle efficiency on a particular drive cycle, but in practice other regional factors could affect the results. For example, the level of humidity will affect HVAC use, and the road conditions (such as terrain, precipitation, and wind) can also affect the efficiency of the vehicle. Our efficiency estimates are based on linear interpolation using test results at three temperatures for each drive cycle. Comparisons with in Yuksel and Michalek [11] suggest that this captures the general shape of the trend reasonably well but coarsely. We also avoid extrapolation beyond the range of temperatures tested and therefore likely make optimistic estimates of vehicle efficiency loss in extreme weather regions.

5.4. Vehicles

We examine only five specific vehicle models for which we have access to laboratory test data at multiple chamber temperatures and multiple drive cycles. Other vehicle models, including more recent model years of the vehicles examined, could have different performance characteristics, temperature sensitivity, etc

5.5. Other externalities

We focus on GHG emissions, but other externalities, including criteria air pollutant emissions and their effect on health, dependence on foreign oil and its relation to energy security and independence, water resource use for energy production, and battery hazardous waste disposal play important roles in guiding policy decisions. In particular, electric vehicle externalities from air pollution may be larger then those for global warming [12, 16, 40].

6. Policy implications

Our results suggest that the GHG-reduction benefits of PEVs have significant regional variability due to grid mix, temperature, and driving conditions as well as differences among vehicle alternatives within each technology class. This suggests that a regionally-targeted vehicle-specific strategy to encourage adoption primarily in areas where specific PEVs provide the largest benefits could increase the GHG reductions achievable under a given budget.

While current federal policy for PEVs is fairly uniform across the US, individual states have adopted differentiated policies including zero-emission vehicle mandates, state tax breaks for PEV purchases, and a range of other incentives, such as subsidized charging infrastructure or access to high-occupancy vehicle lanes for PEV owners. For instance, California, Oregon, New York, New Jersey, Maryland, Connecticut, Rhode Island, Massachusetts, Vermont, and Maine all have policies that mandate sales of vehicles with zero tailpipe emissions (called 'zero emission vehicles' or ZEVs) based on California's policy authorized under section 177 of the Clean Air Act [41]. In urban counties (city driving) of these ZEV states the PEVs we model are lower emitting than the Mazda 3 CV, but they are not all lower emitting in rural counties (highway driving), and some PEVs (e.g.: the Volt) are higher-emitting than the gasoline-powered Prius HEV in all counties of these states.

Further, state subsidies for PEV purchases vary, with the largest subsidies offered in Colorado and, until recently, in West Virginia and Georgia [42], and there is evidence that subsidies increase adoption [43]. West Virginia and Georgia in particular are locations where the GHG case for PEVs in our analysis is less strong, since the gasoline-powered Prius HEV has lower life cycle GHG emissions there than either the Leaf BEV or the Volt PHEV.

Our results suggest that the GHG case for PEVs is generally strongest in urban counties of Texas, Florida, and the Southwestern US followed by New England, and it is generally weakest in the Midwest and the South. However, it is important to note that these estimates are uncertain and dynamic, since (1) the power grid is highly interconnected and changes over the life of the vehicle as the power plant fleet and feedstock prices fluctuate, (2) on-road weather effects on vehicle efficiency may differ from controlled laboratory tests at fixed ambient temperature settings, (3) driving conditions in practice are heterogeneous within each county and are far more diverse than the standard city/highway laboratory tests can capture, and (4) PEV benefits relative to gasoline vehicles vary across different PEV models and depend on which gasoline vehicle the PEV buyer would have purchased if the PEV were not available. The complexity of these uncertain and dynamic regional and vehicle differences makes it difficult to forecast regional GHG benefits of PEVs with certainty, and such challenges pose difficulties for regulators worldwide.

Broadly, regional policies that are more aligned with the GHG benefits we estimate could be more efficient at achieving GHG reductions, though other factors such as regional consumer preferences, political climate, and other externalities also affect regional policy choices. In general, policies that target GHG reductions directly, such as carbon tax or cap-and-trade policies, rather than favoring specific technologies, are likely to be more efficient at achieving GHG reductions, though support for the development and deployment of new technologies can also have dynamic benefits and potentially lead to large long-term benefits if they enable a fleet transition that would not have happened otherwise [44, 45]

Regional differences in GHG emissions from PEVs also have implications for vehicle labeling and regulation. GHG emission estimates used for vehicle fuel economy and environment labels (window stickers) currently report only tailpipe emissions. But upstream GHG emissions from PEV charging can be larger than tailpipe emissions, and they vary regionally. Ideally, future labels will include life cycle emissions estimates that include power plant emissions—but this goal is challenging to achieve with precision given the regional variability and the challenges described previously. Secondly, the US EPA regulates GHG emissions from motor vehicle fleets and currently treats PEVs as though they are zero-emission vehicles when operating on electricity [43]. If future regulations are updated to incorporate upstream PEV emissions from vehicle charging, as they are expected to, regional differences and regional patterns of vehicle adoption will be important to achieving meaningful estimates of GHG emissions from PEVs.

Finally, larger factors can influence policy strategies. For example, when deciding where to allocate scarce public resources, benefits of light-duty transportation electrification must be weighed against benefits that could be achieved in other sectors [46]. Further, our analysis focuses on life cycle emissions directly associated with the vehicles we assess and ignores consequential fleet-wide GHG emission effects of PEV adoption due to alternative fuel vehicle incentives in federal corporate average fuel economy policy and GHG emissions standards. These incentives allow automakers that sell PEVs to meet less-stringent fleet GHG emission standards, at least through 2025, result in net GHG increases when PEVs are sold [47]. This policy effect can be large enough to wipe out any net GHG savings offered by PEV adoption in the near term, although PEV adoption could also have dynamic effects on technology trajectories in the light-duty vehicle fleet that help encourage a long term transition.

Acknowledgments

The authors would like to thank Kevin Stutenberg, Eric Rask, and Forrest Jehlik from Argonne National Laboratory for their help with the dynamometer test data. Funding for this work came from a grant from the Engineering Research and Development for Technology Scholarship Program at the University of the Philippines, a gift from Toyota Motor Corporation, and the Center for Climate and Energy Decision-Making (CEDM), through a cooperative agreement between the National Science Foundation and Carnegie Mellon University (SES-0949710 and SES-1463492). The findings and views expressed are those of the authors and not necessarily those of the sponsors.

References

[1] Samaras C and Meisterling K 2008 Life cycle assessment of greenhouse gas emissions from plug-in hybrid vehicles: implications for policy *Environ. Sci. Technol.* **42** 3170–6

[2] Tamayao M-A M, Michalek J J, Hendrickson C and Azevedo I M L 2015 Regional variability and uncertainty of electric vehicle life cycle CO_2 emissions across the United States *Environ. Sci. Technol.* **49** 8844–55

[3] Yuksel T and Michalek J 2012 Development of a simulation model to analyze the effect of thermal management on battery life *SAE Techical Paper* (doi:10.4271/2012-01-0671)

[4] Anair D and Mahmassani A 2012 *State of Charge: Electric Vehicles' Global Warming Emissions and Fuel-Cost Savings Across the United States* (Cambridge, MA: Union of Concerned Scientists Report) (http://ucsusa.org/sites/default/files/legacy/assets/documents/clean_ vehicles/electric-car-global-warming-emissions-report.pdf)

[5] Graff Zivin J S, Kotchen M J and Mansur E T 2014 Spatial and temporal heterogeneity of marginal emissions: implications for electric cars and other electricity-shifting policies *J. Econ. Behav. Organ.* **107** 248–68

[6] Onat N C, Kucukvar M and Tatari O 2015 Conventional, hybrid, plug-in hybrid or electric vehicles? State-based comparative carbon and energy footprint analysis in the United States *Appl. Energy* **150** 36–49

[7] Archsmith J, Kendall A and Rapson D 2015 From cradle to junkyard: assessing the life cycle greenhouse gas benefits of electric vehicles *Res. Trans. Econ.* **52** 72–90

[8] Karabasoglu O and Michalek J 2013 Influence of driving patterns on life cycle cost and emissions of hybrid and plug-in electric vehicle powertrains *Energy Policy* **60** 445–61

[9] Raykin L, Roorda M J and MacLean H L 2012 Impacts of driving patterns on tank-to-wheel energy use of plug-in hybrid electric vehicles *Transp. Res.* D **17** 243–50

[10] Raykin L, MacLean H L and Roorda M J 2012 Implications of driving patterns on well-to-wheel performance of plug-in hybrid electric vehicles *Environ. Sci. Technol.* **46** 6363–70

[11] Yuksel T and Michalek J J 2015 Effects of regional temperature on electric vehicle efficiency, range, and emissions in the United States *Environ. Sci. Technol.* **49** 3974–80

[12] Holland S P, Mansur E T, Muller N Z and Yates A J 2015 Environmental benefits from driving electric vehicles? *NBER Working Papers* (www.nber.org/papers/w21291)

[13] Yawitz D and Kenward A 2013 *A Roadmap to Climate-Friendly Cars* (Princeton, NJ: Climate-Central) (http://assets.climatecentral.org/pdfs/ClimateFriendlyCarsReport_Final.pdf)

[14] Duvall M, Knipping E and Alexander M 2007 Environmental assessment of plug-in hybrid electric vehicles—volume 1: nationwide greenhouse gas emissions *Electric Power Research Institute Report* (http://epri.com/abstracts/pages/ProductAbstract.aspx?productId=000000000001015325)

[15] Hadley S W and Tsvetkova A A 2009 Potential impacts of plug-in hybrid electric vehicles on regional power generation *Electr. J.* **22** 56–68

[16] Michalek J J, Chester M, Jaramillo P, Samaras C, Shiau C-S N and Lave L B 2011 Valuation of plug-in vehicle life-cycle air emissions and oil displacement benefits *Proc. Natl Acad. Sci. USA* **108** 16554–8

[17] Macpherson N D, Keoleian G A and Kelly J C 2012 Fuel economy and greenhouse gas emissions labeling for plug-in hybrid vehicles from a life cycle perspective *J. Ind. Ecol.* **16** 761–73

[18] Nealer R, Reichmuth D and Anair D 2015 Cleaner cars from cradle to grave: how electric cars beat gasoline cars on lifetime global warming emissions *Union of Concerned Scientists Report* (http://ucsusa.org/sites/default/files/attach/2015/11/Cleaner-Cars-from-Cradle-to-Grave-full-report.pdf)

[19] (Sandy) Thomas C E 2012 US marginal electricity grid mixes and EV greenhouse gas emissions *Int. J. Hydrog. Energy* **37** 19231–40

[20] Siler-Evans K, Azevedo I L and Morgan M G 2012 Marginal emissions factors for the US electricity system *Environ. Sci. Technol.* **46** 4742–8

[21] Alexander M, Knipping E, Yound D, Duvall M, Tonachel L and Hwang R 2015 Environmental assessment of a full electric transportation portfolio—volume 2, greenhouse gas emissions *Electric Power Research Institute Report* (http://epri.com/abstracts/Pages/ProductAbstract.aspx?productId=000000003002006876)

[22] Kelly J C, MacDonald J S and Keoleian G A 2012 Time-dependent plug-in hybrid electric vehicle charging based on national driving patterns and demographics *Appl. Energy* **94** 395–405

[23] Neubauer J 2014 Thru-life impacts of driver aggression, climate, cabin thermal management, and battery thermal management on battery electric vehicle utility *J. Power Sources* **259** 262–75

[24] Kambly K R and Bradley T H 2014 Estimating the HVAC energy consumption of plug-in electric vehicles *J. Power Sources* **259** 117–24

[25] Kambly K and Bradley T H 2015 Geographical and temporal differences in electric vehicle range due to cabin conditioning energy consumption *J. Power Sources* **275** 468–75

[26] Meyer N, Whittal I, Christenson M and Loiselle-Lapointe A 2012 The impact of driving cycle and climate on electrical consumption and range of fully electric passenger vehicles *EVS26–Int. Battery, Hybrid and Fuel Cell Electric Vehicle Symp.* (*Los Angeles*) (www.tc.gc.ca/eng/programs/environment-etv-electric-passenger-vehicles-eng-2904.htm)

[27] Lohse-Busch H, Duoba M, Rask E, Stutenberg K, Gowri V, Slezak L and Anderson D 2013 Ambient temperature (20 °F, 72 °F and 95 °F) impact on fuel and energy consumption for several conventional vehicles, hybrid and plug-in hybrid electric vehicles and battery electric vehicle *SAE Technical Paper* 2013-01-1462 (doi:10.4271/2013-01-1462)

[28] Allen M 2013 Electric Range for the Nissan Leaf and Chevrolet Volt in Cold Weather (http://news.fleetcarma.com/2013/12/16/nissan-leaf-chevrolet-volt-cold-weather-range-loss-electric-vehicle/)

[29] Argonne National Laboratory Downloadable Dynamometer Database (http://transportation.anl.gov/D3/) (accessed 24 April 2015)

[30] United States Environmental Protection Agency Dynamometer Drive Schedules (http://3.epa.gov/nvfel/testing/dynamometer.htm)

[31] Patil R, Adornato B and Filipi Z 2010 Design optimization of a series plug-in hybrid electric vehicle for real-world driving conditions (10-01-0840) *SAE Int. J. Engines* **3** 655–65

[32] US Department of Transportation Federal Highway Administration, NHTS Data Center 2009 *NHTS—Version 2.1 February 2011* (http://nhts.ornl.gov/download.shtml) (accessed 16 April 2014)
[33] National Renewable Energy Laboratory 1991–2005 National Solar Radiation Data Base, Update: TMY3 (http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/tmy3/) (accessed 16 April 2014)
[34] Interpolate Scattered Data—MATLAB Griddata (http://mathworks.com/help/matlab/ref/griddata.html) (accessed 30 January 2015)
[35] Argonne National Laboratory 2014 GREET Life-cycle Assessment Model User's Manual (https://greet.es.anl.gov/publication-greet-model)
[36] United States Environmental Protection Agency 2014 *Inventory of US Greenhouse Gas Emissions and Sinks: 1990–2012* vol 79 (Washington, DC: United States Environmental Protection Agency) (www3.epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2014-Main-Text.pdf)
[37] Venkatesh A, Jaramillo P, Griffin W M and Matthews H S 2011 Uncertainty analysis of life cycle greenhouse gas emissions from petroleum-based fuels and impacts on low carbon fuel policies *Environ. Sci. Technol.* **45** 125–31
[38] Venkatesh A, Jaramillo P, Griffin W M and Matthews H S 2011 Uncertainty in life cycle greenhouse gas emissions from United States natural gas end-uses and its effects on policy *Environ. Sci. Technol.* **45** 8182–9
[39] US EPA OOC eGRID (http://epa.gov/energy/egrid) (accessed 11 February 2016)
[40] Tessum C W, Hill J D and Marshall J D 2014 Life cycle air quality impacts of conventional and alternative light-duty transportation in the United States *Proc. Natl Acad. Sci. USA* **111** 18490–5
[41] United States Environmental Protection Agency 2015 Clean Air Act Text (http://2.epa.gov/clean-air-act-overview/clean-air-act-text) (accessed 6 November 2015)
[42] Alternative Fuels Data Center State Laws and Incentives (http://afdc.energy.gov/laws/state) (accessed 6 November 2015)
[43] Jenn A, Azevedo I L and Ferreira P 2013 The impact of federal incentives on the adoption of hybrid electric vehicles in the United States *Energy Econ.* **40** 936–42
[44] Greene D L, Park S and Liu C 2014 Analyzing the transition to electric drive vehicles in the US *Futures* **58** 34–52
[45] National Resarch Council of the National Academies 2013 *Transitions To Alternative Vehicles And Fuels* (Washington, DC: The National Academies Press) (www.nap.edu/catalog/18264/transitions-to-alternative-vehicles-and-fuels)
[46] Thomas B A and Azevedo I L 2014 Should policy-makers allocate funding to vehicle electrification or end-use energy efficiency as a strategy for climate change mitigation and energy reductions? Rethinking electric utilities efficiency programs *Energy Policy* **67** 28–36
[47] Jenn A, Azevedo I and Michalek J J 2016 Alternative fuel vehicle adoption increases fleet gasoline consumption and greenhouse gas emissions under United States corporate average fuel economy policy and greenhouse gas emissions standards *Environ. Sci. Technol.* **50** 2165–74
[48] US Department of Energy, Office of Energy Efficieny and Renewable Energy, US Environmental Protection Agency Fuel Economy Guide (www.fueleconomy.gov/feg/pdfs/guides/FEG2014.pdf)

Environmental Science & Technology
Policy Analysis
pubs.acs.org/est

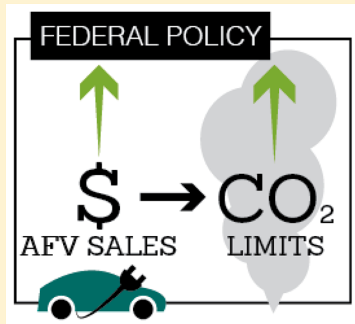
Alternative Fuel Vehicle Adoption Increases Fleet Gasoline Consumption and Greenhouse Gas Emissions under United States Corporate Average Fuel Economy Policy and Greenhouse Gas Emissions Standards

Alan Jenn, Inês M. L. Azevedo, and Jeremy J. Michalek*

Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, United States

Supporting Information

ABSTRACT: The United States Corporate Average Fuel Economy (CAFE) standards and Greenhouse Gas (GHG) Emission standards are designed to reduce petroleum consumption and GHG emissions from light-duty passenger vehicles. They do so by requiring automakers to meet aggregate criteria for fleet fuel efficiency and carbon dioxide (CO_2) emission rates. Several incentives for manufacturers to sell alternative fuel vehicles (AFVs) have been introduced in recent updates of CAFE/GHG policy for vehicles sold from 2012 through 2025 to help encourage a fleet technology transition. These incentives allow automakers that sell AFVs to meet less-stringent fleet efficiency targets, resulting in increased fleet-wide gasoline consumption and emissions. We derive a closed-form expression to quantify these effects. We find that each time an AFV is sold in place of a conventional vehicle, fleet emissions increase by 0 to 60 t of CO_2 and gasoline consumption increases by 0 to 7000 gallons (26,000 L), depending on the AFV and year of sale. Using projections for vehicles sold from 2012 to 2025 from the Energy Information Administration, we estimate that the CAFE/GHG AFV incentives lead to a cumulative increase of 30 to 70 million metric tons of CO_2 and 3 to 8 billion gallons (11 to 30 billion liters) of gasoline consumed over the vehicles' lifetimes − the largest share of which is due to legacy GHG flex-fuel vehicle credits that expire in 2016. These effects may be 30−40% larger in practice than we estimate here due to optimistic laboratory vehicle efficiency tests used in policy compliance calculations.



INTRODUCTION

About 28% of the United States greenhouse gas (GHG) emissions are produced by the transportation sector (the second largest United States GHG source, after the electricity sector), and 62% of these emissions are produced by light-duty vehicles.[1] Light-duty vehicles also consumed 118 billion gallons (450 billion liters) of gasoline in 2012, representing more than half of the petroleum-based fuels consumed in United States transportation.[1] The main United States policy effort to control petroleum consumption and greenhouse gas emissions in the United States light-duty vehicle fleet is the federal Corporate Average Fuel Economy (CAFE) policy and associated Greenhouse Gas Emission standard.

A History of CAFE. In response to the oil crisis of 1973, the United States passed the Energy Policy and Conservation Act of 1975 (Public Law 94163), which included CAFE standards. CAFE mandates that the sales-weighted average fuel efficiency of all new light-duty vehicles sold by each manufacturer in a particular year must meet or exceed a specific target. These targets were initially the same for each manufacturer (although some manufacturers chose to pay fines rather than comply[2]), and separate targets were set for cars and light trucks. The first standards came into effect in 1978 for passenger cars and were followed by standards for light-duty trucks the following year. A timeline of the standards and changes is shown in Figure 1.

The National Highway Traffic Safety Administration (NHTSA) originally promulgated the CAFE standards, but following California's efforts to create state-specific standards and a court ruling in 2007 that required the U.S. Environmental Protection Agency (EPA) to regulate CO_2 emissions as pollutants under the Clear Air Act (Massachusetts versus U.S. Environmental Protection Agency), the rule making for the newest set of CAFE standards and GHG emission standards were passed as a joint set of rules between NHTSA and the EPA in 2010 and came into effect in 2012, applying to model years 2012 to 2016. For the first time, these standards also required carbon dioxide emissions compliance from manufacturers. The EPA regulates fleet average GHG emissions (hereafter referred to as the GHG standard), while NHTSA regulates the corresponding fleet average fuel efficiency (hereafter referred to as the CAFE standard). The NHTSA and EPA standards were harmonized to have comparable stringency,[4,5] but there are also important differences between the two rules.

Received: June 10, 2015
Revised: December 28, 2015
Accepted: January 11, 2016

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

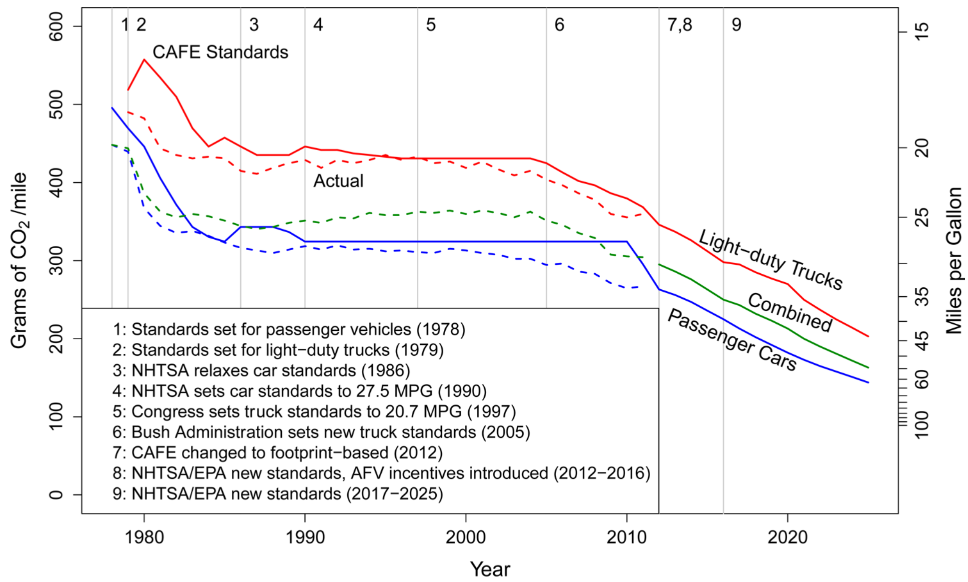


Figure 1. Historical CAFE/GHG Standards and Expected Joint Rule-Making Standard Requirements from 1978 through 2025. Dates correspond to the effective implementation dates of each new policy. Data sources: refs 3−5.

Each agency offers manufacturers compliance flexibility mechanisms that include (1) credits that can be earned if a manufacturer's fleet has lower emissions or higher efficiency than the respective policy requires in a given year and can be traded or used when a manufacturer's fleet would otherwise not comply with the policy, (2) credits for air conditioning improvements, (3) other off-cycle credits for measurable GHG and fuel savings from technologies whose benefits are not measured by the standard laboratory two-cycle test, and (4) incentives for selling AFVs.[5] We focus exclusively on the last effect.

While the two agencies worked in coordination to establish these fuel efficiency and GHG standards, they differ in that the EPA standard allows certain air conditioning improvement credits toward compliance with the GHG standards that NHTSA is not permitted to allow toward compliance with CAFE policy. To address this difference, NHTSA relaxes the stringency of their standard to a level that maintains a harmonized standard with the EPA (see pages 25329−25330 in ref 4), assuming that manufacturers take full advantage of the air conditioning credits (which they are expected to do).

Additionally, NHTSA incentives for AFVs differ from EPA incentives for AFVs due in part to differences in the regulatory authority of the two agencies. The two policies were designed to have comparable stringency, but because they are not identical, it is possible that one standard may be slightly more restrictive than the other for a given manufacturer's fleet in a given year. While it is potentially true that the CAFE standard could be slightly more stringent than the GHG standard for a given manufacturer, the penalty for violating the GHG standard is severe (potential revocation of the license to sell vehicles in the United States), whereas the penalty for violating the CAFE standard is relatively mild ($5.50 per 0.1 mpg violation per vehicle—a quantity that manufacturers have been willing to pay in the past even when standards were far more lax). In particular, the Federal Register notes that "NHTSA recognizes that some manufacturers may use the option to pay civil penalties as a CAFE compliance flexibility—presumably, when paying civil penalties is deemed more cost-effective than applying additional fuel economy-improving technology, or when adding fuel economy-improving technology would fundamentally change the characteristics of the vehicle in ways that the manufacturer believes its target consumers would not accept. NHTSA has no authority under EPCA/EISA to prevent manufacturers from turning to payment of civil penalties if they choose to do so. This is another important difference from EPA's authority under the CAA, which allows EPA to revoke a manufacturer's certificate of conformity that permits it to sell vehicles if EPA determines that the manufacturer is in non-compliance, and does not permit manufacturers to pay fines in lieu of compliance with applicable standards" (ref 5, pp 63130−63131). For this reason, we focus on treating the GHG standard as the binding constraint in our analysis, and we present results for a binding CAFE standard in the Supporting Information.

In addition to changes in average fuel economy targets over time, in 2012, the targets became attribute based; the efficiency target for each vehicle is a function of its footprint (the product of wheelbase and track width—a measure of vehicle size).[4] For both passenger cars and light-duty trucks, vehicles with a larger footprint have less stringent efficiency targets. Each vehicle sold does not necessarily need to comply with the standard associated with its footprint. Instead, the focal year sales-weighted average efficiency of all vehicles sold by each manufacturer must meet or exceed the sales-weighted standard defined by the footprints of the vehicles sold that year (Figure S1, Supporting Information). The intent of the attribute-based standards is to reduce fuel consumption and emissions primarily by encouraging technological improvements across the fleet, rather than shifting consumers into smaller vehicles.[4]

By 2025, the average fuel efficiency of new passenger cars will be required to meet or exceed 54.5 MPG (4.3 L per 100 km) (as measured by a two-cycle laboratory test and based on the EPA GHG standard assuming the entire fleet is able to meet the standard through fuel economy improvements alone).[5] These requirements will likely have strong effects on the vehicle market, both for manufacturers, who must make significant technological improvements to keep pace with the mandate, as well as for consumers, who will have access to a different set of

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Summary of AFV Incentives in the GHG Standard[4,5]

	% VMT on alt fuel, p_j		weighting factor, w_j		multiplier, m_j				
vehicle type	2012−2015	2016−2025	2012−2015	2016−2025	2012−2016	2017−2019	2020	2021	2022−2025
ICV	0	0	1	1	1	1	1	1	1
FFV	50	15	0.15	1	1	1	1	1	1
CNG	100	100	1	1	1	1.6	1.45	1.3	1
BEV	100	100	0	0	1	2.0	1.75	1.5	1
PHEV	29−66	29−66	0	0	1	1.6	1.45	1.3	1
FCV	100	100	0	0	1	2.0	1.75	1.5	1

vehicle options at different prices than they would in the absence of regulation.

The policy will substantially decrease future gasoline consumption and corresponding GHG emissions per mile driven compared to 2009 (see Figure S2 in the Supporting Information for a summary of compliance in 2009 by manufacturer). Broadly speaking, for both cars and trucks, the American manufacturers have historically tended to treat the CAFE standard as a binding constraint, while Asian manufacturers tended to overcomply and European manufacturers tended to undercomply (and therefore paid penalties). As the standard increases in stringency, and as penalties for violation are increased, manufacturers will need to implement vehicle design changes and/or shift the portfolio of vehicles they sell in order to comply. Since penalties for violation of the new GHG standards are higher than those of the older CAFE standard, we follow prior analysis[6] in assuming the standards will be binding for all manufacturers in the future (with the exception of Tesla—a unique automaker focused on low volume electric vehicles). Figure S2 in the Supporting Information shows that in the 2009 fleet no automaker other than Tesla would have satisfied the 2016 standards, providing further evidence that the standards are binding. However, if any firms were to find the CAFE/GHG standard to be nonbinding without the AFV incentives, the policy and the incentives would be irrelevant for that manufacturer.

The Congressional Budget Office (CBO) noted in a 2012 report "With CAFE standards in place···putting more electric (or other high-fuel-economy) vehicles on the road will produce little or no net reduction in total gasoline consumption and greenhouse gas emissions".[6] This is because future stringent GHG standards are expected to be binding with high penalties for violation, and under a binding standard, the annual target would be achieved regardless of whether AFVs are sold. This effect—where efforts to reduce emissions in one area lead to increased emissions elsewhere, resulting in no net benefit—has been referred to as "leakage". Goulder et al.[7] also note this leakage effect in relation to state Pavley limits on vehicle greenhouse gas emissions. Leakage is not a property of the CAFE/GHG policy itself but rather a description of the fleet-wide implications of other policies intended to reduce emissions or gasoline consumption in a particular subset of the United States fleet when implemented in the presence of binding national standards.

We find that this leakage effect is now amplified by AFV incentives in CAFE/GHG standards. Beginning in 2012, the EPA/NHTSA policy includes incentives that encourage automakers to produce AFVs by allowing automakers that sell AFVs to meet less-stringent fleet standards. The rules offer different incentives for flex fuel vehicles (FFVs), compressed natural gas vehicles (CNG), battery electric vehicles (BEVs), plug-in hybrid electric vehicles (PHEVs), and fuel cell vehicles (FCVs). There are two types of AFV incentives in the GHG standard: weighting factors and multipliers. A weighting factor reduces the effective emissions rate for AFVs used in compliance calculations, allowing AFVs to count as though they have lower emissions than they actually do and relaxing the stringency of the automaker's fleet standard. A multiplier allows each AFV sold to count as more than one vehicle sold in compliance calculations, further relaxing stringency of the automaker's standard (whenever the AFV is lower emitting than the manufacturer's average vehicle). Table 1 summarizes the weights and multipliers in the GHG policy from 2012 to 2025. We estimate the magnitude of the resulting implications of AFV incentives in a binding GHG standard for fleet gasoline consumption and greenhouse gas emissions. The EPA also notes this effect and estimates the decrease in GHG emission reductions due to projected PHEV and BEV adoption in model years 2017 to 2025 under these incentives (ref 5, pp. 62811, ref 18 p4−141). They argue that "EPA believes it is worthwhile to forego modest additional emissions reductions in the near term in order to lay the foundation for the potential for much larger 'game-changing' GHG emissions and oil reductions in the longer term." The Supporting Information provides additional estimates for the case when the CAFE standard (which has statutory weighting factors for AFVs but not multipliers) is binding.

Literature Review. The CAFE policy has had a profound impact on transportation in the United States; over the last several decades, it has affected the emissions of hundreds of millions of vehicles and reduced consumption of gasoline on the order of billions of gallons, as the following studies indicate. The effectiveness and efficiency of the CAFE policy for reducing emissions and oil consumption has been well studied—and hotly debated. In a 1998 evaluation of CAFE standards, Greene[8] argued that fuel economy regulation has been economically efficient and, despite a potential rebound effect, has saved consumers $50 billion annually (Azevedo[9] estimates that direct rebound effects in personal transportation likely range from 4% to 87%; however, recent studies suggest that the short-term price elasticity for fuel, used as a proxy for direct rebound effects, is fairly inelastic, and there is some indication that it has been decreasing over time.[10])

Greene also warns that "simply because a corporate average fuel economy formula worked well in the past does not mean that a more efficient formulation does not exist". Indeed, most economists argue that imposing gasoline taxes can achieve the same outcomes as CAFE more efficiently—though implementation of fuel taxes is controversial and politically challenging. For example, Kleit[11] reports that a gas tax of $0.11 per gallon would lead to the same gasoline savings as the CAFE standards, while costing far less (a $4 billion welfare loss due to CAFE compared to a $290 million welfare cost due to gasoline taxes). Similarly, Austin and Dinan[12] use a Bertrand equilibrium model

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

to project responses to fuel efficiency standards and find that gasoline taxes would result in around 60% lower welfare losses while achieving the same oil consumption decrease. However, Gerard and Lave[13] argue that such taxes ought to supplement existing CAFE standards, rather than replace them, because CAFE inefficiencies are mitigated with gas taxes that internalize externalities and because consumers use higher implicit discount rates than social discount rates, and they tend to purchase less-efficient vehicles among those with equivalent lifetime costs.

A range of studies have followed the announcement and implementation of the 2012–2016 CAFE/GHG standards, estimating fuel and emissions savings using economic equilibrium models,[14,15] life-cycle assessment,[16] and decision theory.[17] The EPA also released a report evaluating the effect of the 2012–2016 standards,[18] estimating 1 billion metric tons of CO_2 reductions and savings of 1.8 billion barrels of oil over the lifetime of new vehicles sold during the period. By 2050, the EPA expects that CAFE standards will lead to reductions of 500 million metric tons of CO_2 annually.[18] The United States emissions from the transportation sector are currently about 1.8 billion metric tons of CO_2 annually,[19] so this is a substantial reduction in emissions. CAFE policy achieves these reductions by incentivizing automakers to redesign vehicles, implement fuel savings technologies, and adjust fleet sales mix (e.g., via strategic pricing). Whitefoot et al.[20] argue that firms may rely primarily on vehicle design changes rather than strategic pricing to comply with standards, although Shiau et al.[21] suggest that the CAFE policy can be ineffective at causing changes to vehicle design when the standard is set too high without a corresponding increase in the penalty for violation. Whitefoot and Skerlos[22] argue that footprint-based standards incentivize automakers to increase vehicle size, potentially undermining fuel economy gains by an estimated 1 to 4 MPG and increasing new vehicle emissions by 5% to 15%.

AFV incentives in CAFE policy further complicate the policy's effects. Anderson and Sallee[23] estimate that the ability of automakers to exploit flex-fuel vehicle incentives reduces CAFE compliance costs dramatically. Goulder et al.[7] show that because of federal CAFE standards, the California Zero Emission Vehicle (ZEV) regulation has no net effect on fuel consumption or emissions due to the leakage effect; sales of fuel efficient vehicles in California and other ZEV states are balanced by sales of less-efficient vehicles in other states, resulting in no net benefits at the national level.

However, because of AFV incentives in CAFE/GHG policy this leakage effect is compounded and sale of AFVs results in increases of fleet emissions and fuel consumption. EPA estimates the net effect of the incentives for BEVs and PHEVs in the GHG standard on fleet GHG emissions to be an increase of 56 to 101 million metric tons of CO_2 equivalent for model year 2017–2025 based in part on detailed models of the most cost-effective ways industry is expected to meet the standards (ref 5 p62811, ref 18 p4–141). We perform an independent assessment of the effect for all AFVs; we derive a closed form expression for the change in fleet emissions and gasoline consumption per AFV sold for the period 2012 through 2025; and we estimate the net effect using a range of sales projections.

■ DATA AND METHODS

GHG Standards and AFV Incentives. We assume that there will be no changes in the policy design between now and 2025, that the total number of vehicles sold by each manufacturer is not affected by the AFV incentives, and that the GHG standards are binding (i.e., we assume that each manufacturer will comply with future GHG standards without significantly exceeding them). Both the EPA (ref 4, pp 25342–25343) and the Congressional Budget Office[6] make similar assumptions in their analysis of the effects of the CAFE/GHG standards. When a manufacturer complies exactly with the GHG standards, it satisfies the following equation:

$$\frac{\sum_{j\in J} n_j s_j}{N} = \frac{\sum_{j\in J} n_j r_j}{N} \tag{1}$$

where n_j is the number of units of vehicle model j sold by the manufacturer in the focal year, s_j is the footprint-based GHG standard associated with vehicle model j in the focal year, r_j is the GHG tailpipe emission rate for vehicle model j, $N = \sum_{j\in J} n_j$ is the total number of vehicles sold by the manufacturer in the focal year, and J is the set of all vehicle models offered by the manufacturer. EPA policy requires the sales-weighted average emission rate to be less than or equal to the standard. We assume the standard is binding (an active constraint), and thus eq 1 enforces an equality.

However, eq 1 does not account for the fact that the GHG standard incorporates a set of AFV incentives. To account for AFV incentives, we partition the set of vehicle models J into the subset of conventional vehicles, J_C, and the subset of alternative fuel vehicles, J_A. The GHG policy includes weighting factors, w, that reduce the effective emission rate attributed to AFVs in compliance calculations, allowing AFVs to count as though they have lower emissions than they actually do. This effectively relaxes the standard. A multiplier, m, allows each AFV sold to count as more than one vehicle sold in compliance calculations and can either decrease or increase the stringency of the standard depending on whether the AFV is lower or higher emitting than the manufacturer's average vehicle, respectively. The resulting relation for the GHG standard with AFV weights and multipliers is

$$\frac{\sum_{j\in J} n_j s_j}{N} = \frac{\sum_{j\in J_C} n_j r_j + \sum_{j\in J_A} n_j m_j (w_j p_j r_j^{A} + (1-p_j) r_j^{G})}{\sum_{j\in J_C} n_j + \sum_{j\in J_A} n_j m_j} \tag{2}$$

where $w_j \in [0, 1]$ is the weighting factor for AFV model j, $m_j \geq 1$ is the multiplier for AFV model j, r_j^{A} is the emission rate of AFV model j when operating on its alternative fuel (including some upstream emissions, such as power plant emissions for charging BEVs or PHEVs), r_j^{G} is the tailpipe emission rate of dual-fuel AFV model j when operating on gasoline, and p_j is the assumed portion of AFV miles propelled using the alternative fuel ($p_j = 1$ for pure AFVs but $p \in (0, 1)$ for dual fuel vehicles that use a mix of gasoline and an alternative fuel, such as FFVs and PHEVs). Note that in the EPA rule, because r and r_j^{G} historically measure only tailpipe emissions and ignore upstream emissions from gasoline production and distribution supply chains, and because differences in upstream emissions are important when comparing AFVs to gasoline vehicles, the estimates of AFV emissions r_j^{A} used in compliance calculations are modified to estimate *relative* emissions differences. Specifically, upstream emissions for the average gasoline vehicle are subtracted from the overall estimate of tailpipe + upstream AFV emissions to produce a relative AFV emission rate estimate r_j^{A} (see p 62822 of refs 4 and 5). Table 1 summarizes

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX–XXX

weights, multipliers, and the portion of vehicle miles traveled (VMT) operating on the alternative fuel assumed by the EPA for each of the AFV types included in the 2012–2016 and 2017–2025 rules.

For the particular case when there is no change in the manufacturer's GHG target (e.g., no change in vehicle footprint) induced by the AFV incentives, the net change in GHG emissions associated with vehicle operation, $\Delta\gamma$, is

$$\Delta\gamma = v\left(\sum_{j\in J_A} n_j'((1 - m_j w_j)p_j r_j^A + (m_j - 1)(\bar{s}' - (1 - p_j)r_j^G))\right) \quad (3)$$

where v is the assumed lifetime vehicle miles traveled for all vehicles, n_j' is the sales volume of vehicle model j given the AFV incentives, and $\bar{s}'$ is the manufacturer's sales-weighted GHG target, given the sales mix under the AFV incentives (see the Supporting Information for derivation and for the general case). Examining partial derivatives reveals that net GHG emissions increase as weighting factors, w, are reduced. Net GHG emissions also increase as dual-fuel AFV's gasoline emission rates, r^G, are reduced (holding other factors constant). If an AFV has lower weighted emissions than the manufacturer's GHG standard, then net GHG emissions increase as the multiplier, m, increases and as the AFV sales volume, n,′ increases. The effect of other factors, p and r^A, depends on the values of w and m. When the multiplier is 1 and the weighting factor is 1, the AFV incentive effect is zero. For $m > 1$ or $0 \le w < 1$, the effect of AFV incentives is to increase net emissions (whenever AFVs are lower emitting than the fleet average).

Similarly, we can determine the net gasoline consumption change, $\Delta\lambda$, as a result of the GHG policy:

$$\Delta\lambda = v\delta\left(\sum_{j\in J_A} n_j'\left(\left(\frac{n_j}{n_j'} - m_j w_j\right)p_j r_j^A + (m_j - 1)(\bar{s}' - (1 - p_j)r_j^G)\right)\right) \quad (4)$$

where δ = 1 gallon of gasoline/8887 g of CO_2 is the reciprocal of the carbon dioxide emissions produced per gallon of gasoline combusted (refer to the Supporting Information for derivation). The change in gasoline consumption due to the GHG policy is proportional to the change in emissions if the AFV incentives do not induce additional AFV sales ($n_j' = n_j\ \forall_j \in J_A$).

Net Effects of AFV Incentives for Vehicles Sold between 2012 and 2025. To estimate the net effect of AFV incentives on fleet tailpipe and power plant emissions associated with vehicle operation (i.e., ignoring differences in vehicle manufacturing emissions or end of life emissions for AFVs), we apply projections of AFV sales through 2025 from the reference case scenarios of the EIA's Annual Energy Outlook (AEO) reports in 2012 through 2015 (Figure S5, Supporting Information). We compare four different AEO projections because the sales of AFVs, particularly FFVs, are substantially higher in the 2012 projections (at nearly 1 million sales annually) but have since been adjusted downward in the 2013 projections before increasing in the 2014 and 2015 projections.[24−27]

The AEO reports provide projections of sales for $PHEV_{10}$, $PHEV_{40}$, BEV_{100}, and FFVs.[24−27] We select representative vehicles in each vehicle technology category: The Toyota Prius PHEV, Chevrolet Volt, and Nissan Leaf are used as proxies for the AEO's $PHEV_{10}$, $PHEV_{40}$, and BEV_{100}, respectively. For the representative FFVs, we draw from historical sales-weighted emissions rates, r^A and r^G, of FFVs over the past decade. Estimates may vary for AFVs in other classes (e.g., SUVs, trucks, etc.).

As a base case, we track the net change in GHG emissions, $\Delta\Gamma_t$, annually (where t = {1, 2, ..., 26} refer to years {2012, 2013,..., 2037}, respectively) using United States average estimates of annual VMT as a function of vehicle age based on NHTS survey data[28] summarized in Table S2 of the Supporting Information ($v = \sum_{\tau=1}^{L} v_\tau \approx$ 157,000 mi). We assume each vehicle has a lifetime of L = 12 years.

Again assuming the AFV incentives do not cause a change in the manufacturer's GHG target (e.g., no change to vehicle footprint—see Supporting Information for the general case), the net change in emissions during year t due to vehicles sold in years τ = {1, ..., t} is computed as

$$\Delta\Gamma_t = \sum_{\tau=1}^{t}\sum_{j\in J_A} v_{t-\tau} n_{j\tau}'((1 - m_{j\tau}w_{j\tau})p_{j\tau} r_{j\tau}^A + (m_{j\tau} - 1)(\bar{s}_\tau' - (1 - p_{j\tau})r_{j\tau}^G)) \quad (5)$$

We account for the cumulative change in emissions due to vehicles sold from 2012 to 2025 due to the AFV incentives, but because emissions from these vehicles are produced in years following the vehicle sale, we account for cumulative emissions through 2037 ($\sum_{t=1}^{26}\Delta\Gamma_t$, where $w_{j\tau} = m_{j\tau} = 1\ \forall\tau > 14$; $v_{t-\tau} = 0\ \forall(t - \tau) > L$). We compute gasoline consumption implications in a similar way, but because we lack counterfactual projections of AFV sales in the presence versus absence of the incentives, we focus on the case where AFV sales are unchanged by the incentives and leave alternative scenarios for future work given the uncertainty and the complexity of interactions between incentive-induced sales, weights, and multipliers.

Table S1 of the Supporting Information summarizes emission rates for a set of United States AFVs based on EPA estimates measured via the two-cycle tests used in CAFE/GHG compliance calculations.[29] Emissions associated with electricity consumption are also from EPA estimates; we adopt their figures for upstream electricity GHG emission factors (conversion to emission rates from Wh per 100 mi by EPA methods outlined on page 62822 of ref 5). In the sensitivity analysis, we test the importance of this assumption. The EPA currently considers BEV emissions and PHEV emissions while operating on electricity to be 0 g of CO_2 per mile in compliance calculations. Values for the proportion of VMT, p, propelled by the alternative fuel are also taken from EPA estimates (Table 1).[4,5]

Sensitivity Analysis. The two-cycle test used for measuring CAFE/GHG compliance is known to produce optimistic estimates relative to typical on-road driving patterns.[33] The fuel economy displayed on current vehicle window stickers instead reports the newer five-cycle based testing, and the EPA uses 5-cycle measurements in regulatory impact analysis.[18] If real-world on-road emissions (estimated using the five-cycle test rates), r_5, are φ times as large as two-cycle test emission rates, r, for all vehicles, so that $r_{5j}^A = \varphi r_j^A$ and $r_{5j}^G = \varphi r_j^G\ \forall j \in J$, then the on-road emissions effect of the AFV incentives increases by a factor of φ. These factors are summarized in Table S1 of the Supporting Information.

The second assumption we examine is the grid emissions used in the charging of electric vehicles. The EPA method uses projections of 2030 national average of projected marginal grid emission rates, and we compare this to estimates of the emissions over ranges of recent regional marginal grid emission rates.[25] We adopt a base case "mid" scenario using the EPA

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

projected emission factor and estimate the upper and lower ranges of grid emissions using the lowest and highest annually averaged marginal emission rates by North American Electric Reliability Corporation (NERC) regions from 2007 as estimated by Siler-Evans et al.[30] These range from 530 to 790 kg/MWh. Average emission rates for smaller grid regions ranging from 300 to 1000 kg/MWh have also been used in electric vehicle studies,[33] but given the consequential framing of our analysis, we focus on marginal emission factors, which estimate the effect of changes in the system that result from new electricity demand.

Finally, in the Supporting Information, we examine the case where the CAFE standard is binding instead of the GHG standard.

■ RESULTS

We start by showing the effect of the weights and multipliers for one specific AFV. Figure 2 illustrates how the inclusion of

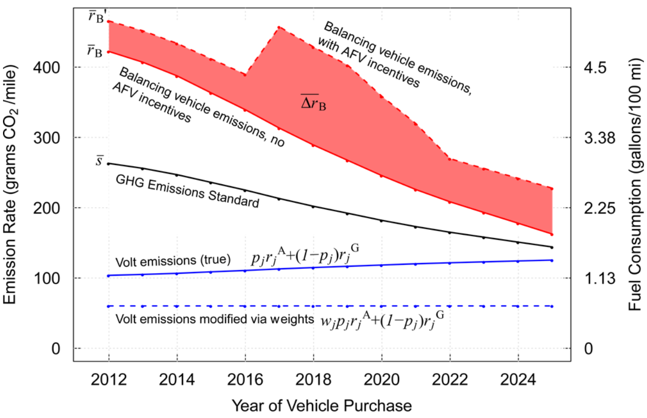


Figure 2. Illustration of emission rates for a Chevrolet Volt and its balancing vehicle (shown here for the case of equal sales volume and no sales induced by the AFV incentive). The balancing vehicle is the vehicle whose emission rate, when averaged with the Volt emission rate using the GHG compliance formula, results in satisfying the GHG standard exactly—shown both with and without AFV incentives. The shaded area represents the increase in average balancing vehicle emission rate due to AFV incentives. Equations are described in the Data and Methods section.

GHG AFV incentives results in increased emission rates for a Chevrolet Volt. The black line shows the annual GHG emissions standards with which the manufacturer needs to comply. If one vehicle has emissions lower than the standard, a second "balancing vehicle" can be sold with higher emissions such that the average emission rate of the two vehicles is equal to the standard. This is an illustrative case with a single balancing vehicle model, equal sales volume for the AFV and its balancing vehicle, and no change in sales volume induced by the AFV incentives. Without AFV incentives, the average of the Volt emission rate (solid blue) and the balancing vehicle emission rate (solid red) is equal to the standard with which the manufacturer would need to comply in each year. The Volt emissions appear to increase over time only because the EPA uses AFV upstream emission estimates *relative* to the upstream emissions of an average conventional internal combustion vehicle (as described earlier), which decrease over time as the standards become more stringent (see p 62822 of ref 4). With the AFV incentives, the adjusted emission rate for the Volt used in GHG accounting calculations is artificially lowered using a weighting factor (dotted blue). The balancing vehicle (dotted red) produces higher emissions for two reasons; between 2012 and 2016, the weighting factor allows the balancing vehicle to be a higher-emitting vehicle, and after 2016, the inclusion of a multiplier, m, greater than one compounds this effect. The net increase in the average emission rate resulting from the AFV incentives is the difference between the red lines (shaded area in Figure 2). For the Volt, this increase ranges from ~40 gCO_2/mi (25 g/km) in 2012–2016 to 140 gCO_2/mi (87 g/km) in 2017. We perform a similar assessment for the AFVs listed in Table S3 in the Supporting Information and find that the increase in emissions ranges between 10 and 400 gCO_2/mi (6 to 250 g/km)—a range comparable to the emissions that would have been created if an extra conventional light-duty vehicle's emissions were added to the fleet's emissions each time an AFV is sold in place of a conventional vehicle (a Toyota Camry is 330 gCO_2/mi (200 g/km)).

The net lifetime increase in fleet GHG emissions and gasoline consumption for several AFVs is shown in Figure 3

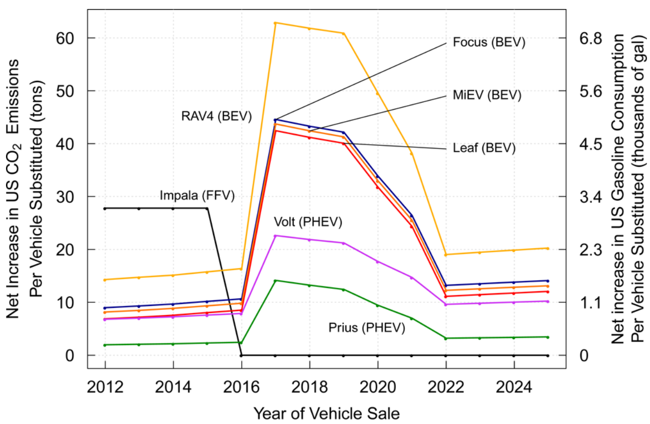


Figure 3. Change in fleet GHG emissions and gasoline consumption each time an AFV is sold in place of a conventional vehicle due to AFV incentives under a binding GHG standard (shown here assuming no change in the manufacturer's footprint-based GHG standard induced by the incentives).

(again for the case of no change to the manufacturer's GHG target induced by the incentives). The greatest increase occurs for battery electric vehicles (BEVs), such as the Nissan Leaf and the Ford Focus BEV, because AFV incentives for these vehicles have weighting factors of $w = 0$ and multipliers as high as $m = 2$. The Chevrolet Volt and Toyota Prius PHEV follow a similar pattern at lower magnitude. Flex fuel vehicles benefit from a 0.15 weighting factor and assumed 50% of VMT propelled by ethanol, both of which expire in 2016.

We also estimate the cumulative increase in GHG emissions resulting from AFV incentives from 2012 to 2025. We use the AEO vehicle sales projections made in 2012, 2013, 2014, and 2015 reports, as explained in the Data and Methods section.[24−27] The results are shown in Figure S3 in the Supporting Information. The largest source of emissions difference between vehicle technologies is caused by the difference in projected sales from the AEO reports. The FFVs have the highest sales in both cases and as a result produce the highest cumulative increase in emissions, although the emissions from FFVs peak earlier, as their AFV incentives expire first. Despite relatively large differences in projected sales of plug-in electric vehicles, we find that the cumulative

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

emissions effect is comparable across technologies from sales in 2012 through 2025, ranging between 2 and 11 million metric tons of increase in CO_2 emissions for each technology using 2013 projections. The net effect of the AFV incentives is an increase of 30 to 70 million metric tons of CO_2 emitted over the lifetime of the vehicles sold during this period. This is the equivalent of relaxing the GHG standard by about 0.8−1.5% (assuming no change in total sales). The effect of AFV incentives on gasoline consumption depends on the change in AFV sales induced by the incentives. Assuming no change in sales, the incentives result in 3−8 billion gallons (11−30 billion liters) of gasoline consumed over the lifetime of the vehicles sold during this period.

Sensitivity Analysis Results. We calculate the difference in emissions between two-cycle tests (used to measure fuel economy for compliance calculations) and five-cycle tests (used to measure fuel economy for vehicle window stickers), which provide more accurate estimates of on-road vehicle fuel economy[34] (Table S1, Supporting Information). Emissions estimates from the five-cycle test are 1.3 to 1.4 times as large as those from the two-cycle test for the vehicle models we examine, suggesting (if the ratio were comparable for all vehicle models) that the on-road emissions implications of the AFV incentives could be 30−40% higher than our base estimates made using CAFE/GHG 2-cycle tests.

Due to uncertainty in emissions from the electric grid resulting from charging of BEVs and PHEVs (refer to Table S3 in the Supporting Information for efficiency of BEVs and PHEVs), we also compare the EPA's projection of incremental grid emission factors in 2030 against estimated marginal emissions rates of different NERC regions in 2007.[30] We use the low-emitting Western Electricity Coordinating Council (WECC) region as a low case and the high-emitting Midwest Reliability Organization (MRO) region as a high case. As shown in Figure 4, the emissions from the EPA projected national grid emissions is closer to the low case, but we find

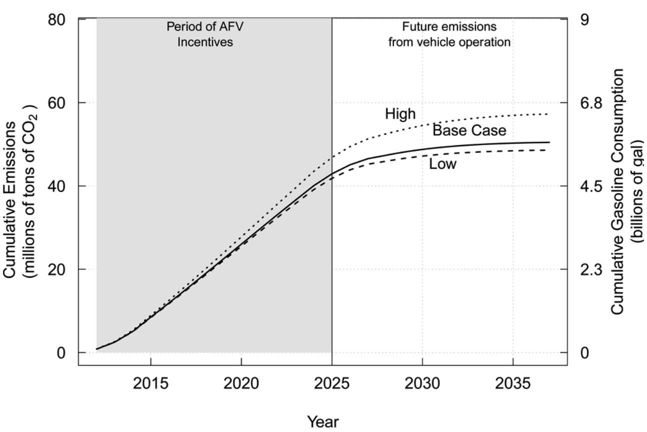


Figure 4. Increase in cumulative emissions due to AFV incentives based on EIA AEO 2015 Alternative Vehicle Sales Forecasts under a binding GHG standard (shown here assuming no change in the manufacturer's footprint-based GHG standard and no change in AFV sales induced by the incentives). High scenario: highest recent marginal emission rate in the United States by NERC region (MRO, Midwest at 786 kg CO_2/MWh). Base case scenario: EPA projected national average incremental emission rate in 2030 (base case: 534 kg CO_2/MWh). Low scenario: lowest recent marginal emission rate in the United States by NERC region (WECC, West at 464 kg CO_2/MWh).

that the total emissions vary by less than 30% from the lowest and highest estimates of 28 to 38 million tCO_2, respectively. Currently, plug-in electric vehicle adoption is concentrated in regions that have lower marginal emission rates.[35]

In the Supporting Information, we also develop a similar analysis for the case where the CAFE standard is binding rather than the GHG standard. We find that the emissions consequences per AFV sold do not peak in 2017 (Figure S6) under a binding CAFE standard as they do under a binding GHG standard (Figure 3) because the CAFE standard has no AFV multipliers. However, the overall cumulative emissions implications of the AFV incentives are comparable under a binding CAFE standard to our estimates under a binding GHG standard (see Supporting Information for details).

Additionally, we ignore the effects of other flexibility mechanisms in CAFE/GHG policy, such as off-cycle credits and credit trading. These credits could interact with the AFV incentives we analyze. For example, if the credits effectively loosen the GHG standard observed by automakers, then the resulting effective $\bar{s}$ in eqs 3−5) may increase, resulting in larger emissions implications than we estimate here for years with multipliers greater than one. We leave analysis of other flexibility mechanisms for future work.

DISCUSSION

We estimate net increases in GHG emissions and gasoline consumption as a result of AFV incentives in a binding light-duty vehicle GHG policy under the assumption that the GHG policy may affect vehicle design and sales mix but not total vehicle sales. We find under fairly general conditions that reducing AFV weighting factors results in increased fleet emissions and gasoline consumption. Increasing AFV multiplier factors also results in increased emissions and gasoline consumption when the manufacturer's incentive-weighted AFV emissions are lower than its fleet average. Further, and counterintuitively, increased sales of AFVs in place of conventional vehicles results in increased United States fleet emissions and gasoline consumption because of the incentives. Fleet-wide gasoline consumption also increases as any dual-fuel AFV technology's gasoline consumption rate is reduced (holding all other factors constant). These outcomes are further modified if the AFV incentives induce a change in the manufacturer's sales mix that significantly affects its GHG target (e.g.: a change in the size of the vehicles sold), and any change in vehicle miles traveled, such as a rebound effect induced by lower operation costs or reduced travel due to electric vehicle range limitations, could further modify fleet-wide implications.

Using sales projections from the AEO 2012−2015 reports,[24−27] we estimate the net effect of the AFV incentives in the GHG standard from vehicles sold from 2012 to 2025 (assuming a 12 year life) is an increase of 30 to 70 million metric tons of CO_2 (50% to 75% due to FFVs) relative to the same policy without AFV incentives (or, equivalently, relative to the same policy if there are no AFV sales). Gasoline consumption implications depend on AFV sales induced by the incentive, but assuming no induced sales implies 3.4 to 7.9 billion additional gallons (11 to 30 billion liters) of gasoline consumed. On-road effects may be 30−40% higher in practice, since our base case analysis is based on optimistic 2-cycle laboratory tests used in CAFE/GHG compliance calculations. Therefore, we estimate the on-road effect as about 40 to 100 million metric tons of CO_2. For comparison, EPA estimates a similar range of emissions (56 to 101 million metric tons of

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

CO_2) for a narrower set of technologies (BEVs and PHEVs) in a shorter period (2017−2025). The difference is due in part to EPA using more optimistic projections of plug-in electric vehicle sales than EIA projections. Our estimates represent about 1−2% of total estimated GHG savings from CAFE/GHG policy, and the net effect on fleet-wide GHG emissions is approximately equivalent to relaxing the overall GHG standards by 0.8% to 1.5%. The policy also has implications for other air pollutants not examined here, which could have large social costs.[36,37]

■ POLICY OPTIONS

The fleet-wide effects we identify under binding GHG standards occur as a result of the interaction of AFV incentives in the GHG policy with an increase in AFV sales, driven largely by state policies. Candidate approaches to addressing this issue might include (1) making no policy changes, (2) eliminating the AFV incentives, (3) eliminating policies that encourage AFV sales, (4) redesigning policies, or (5) considering alternative policies. We examine each approach in turn:

(1) **No Policy Change:** Tolerating the near term emissions and gasoline consumption increases we identify in pursuit of long-term reductions is an option, since the long run emissions and gasoline savings of a transition to AFVs are likely to more than compensate for the short-term increases we estimate, and AFV implementation efforts may further generate positive network externalities.[5,38,39] But future benefits attributed to these policies are only realized if the policies in question succeed in securing a transition to AFVs that would not have happened otherwise. Or, if such policies accelerate a transition that would have happened more slowly otherwise, the benefits of the policy are those associated with the change in the transition interval enabled by the policy. Depending on the magnitude of the policy's effect in accelerating a transition, the long-term benefits of the policy may or may not outweigh the near term increases in emissions and gasoline consumption we estimate.

(2) **Eliminate AFV Incentives:** Eliminating the CAFE/GHG AFV incentives would eliminate the increase in fleet emissions per AFV sold but not the emissions leakage effect (i.e., AFV adoption would produce no net change in fleet emissions or gasoline consumption), and the resulting standards may be more difficult and expensive for automakers to achieve, given low gas prices and consumer preferences for large performance vehicles. In fact, the negotiations in setting policy for the CAFE/GHG standard may have resulted in less stringent fuel efficiency and GHG emissions targets had the incentives been excluded.

(3) **Eliminate Policies That Encourage AFV Sales:** Our analysis shows that reducing AFV sales (e.g., by eliminating policies that encourage or mandate AFV adoption) through 2025 would reduce short-term fleet emissions and gasoline consumption. However, such an option could stall efforts to put the fleet on a path to transition that would take over in a decade even if the ideal technology and infrastructure were available at competitive costs today.

(4) **Redesign Policies:** Improved coordination of federal and state policy design could potentially help to reduce negative interactions among policies because the fleet-wide emissions and gasoline consumption effects we estimate are proportional to the number of AFVs sold, and the state zero-emission vehicle policy represents the largest effort to increase the number of AFVs sold. But coordination is nontrivial; the new CAFE/GHG standards themselves were created as a federal compromise with California, which wanted more stringent state standards.

(5) **Alternative Policies:** Pricing externalities at a value equal to the estimated marginal damage caused to society is among the most efficient options for achieving end goals, but public support for such policies is low in the United States, even if tax revenues are returned to American households.[40] Alternative policies such as regulating CO_2 as a pollutant, subsidizing fuel-efficient vehicles, and requiring high fuel efficiency, are more politically palatable. Nevertheless, continued attempts to persuade the public and lawmakers of the benefits of an efficient externality pricing approach that addresses end goals directly, rather than favoring specific technologies, remains important. While higher prices on gasoline, electricity, and other fuels to reflect the damages they cause are not the only mechanisms needed to secure a transition to alternative fuel vehicles or to manage climate change and air pollution, they would help to mitigate some of the key unintended and often difficult-to-spot effects of interactions among well-intentioned policies.

With the current federal CAFE/GHG policy in place, other federal and state policies that increase AFV market share will result in increased fleet-wide United States greenhouse gas emissions and gasoline consumption through at least 2025. It is hoped that understanding this effect can inform future federal and state policy design while also informing policymakers in other regions with related automotive policies, such as China and the European Union, of the effects of interactions between fleet standards and mechanisms that encourage adoption of specific technologies.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.5b02842.

Additional detail on the derivation of equations for a binding GHG standard, derivation of equations for a binding CAFE standard, results for the binding CAFE case, and additional figures and tables providing additional information about the attribute-based CAFE/GHG standards and their stringency with respect to recent automaker vehicle fleets, detail on projected cumulative emissions based on several AEO vehicle sales projections, comparisons of two-cycle versus five-cycle vehicle efficiency measurements, data on declining annual VMT over a vehicle's life, and a list of AFV attributes used by the EPA. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (412) 268-3765. E-mail: jmichalek@cmu.edu.

Notes

The authors declare no competing financial interest.

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

■ ACKNOWLEDGMENTS

We thank Roberts French and Jeff Alson from the U.S. EPA, John Whitefoot from NHTSA, Rick Gazelle and William Chernicoff from Toyota Motor North America, and Andrew Yates from the University of North Carolina, Chapel Hill for their assistance in understanding CAFE and GHG Emission standards regulations. This work was funded in part by the Center for Climate and Energy Decision Making (SES-0949710 and SES-1463492) through a cooperative agreement between the National Science Foundation and Carnegie Mellon University and by Toyota Motor Corporation. The views expressed are those of the authors and not necessarily those of the sponsors.

■ ABBREVIATIONS

CAFE, corporate average fuel economy; GHG, greenhouse gas; AFV, alternative fuel vehicle; NHTSA, National Highway Traffic Safety Administration; EPA, United States Environmental Protection Agency; CBO, Congressional Budget Office; FFV, flex-fuel vehicle; CNG, compressed natural gas; BEV, battery electric vehicle; PHEV, plug-in hybrid electric vehicle; FCV, fuel cell vehicle; ZEV, zero emission vehicle; AEO, annual energy outlook; VMT, vehicle miles traveled; EIA, United States Energy Information Agency; NERC, North American Electric Reliability Corporation

■ REFERENCES

(1) Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990−2013. U.S. Environmental Protection Agency, 2015. http://www3.epa.gov/climatechange/Downloads/ghgemissions/US-GHG-Inventory-2015-Main-Text.pdf (accessed December 8, 2015).

(2) Summary of CAFE Fines Collected. National Highway Traffic Safety Administration, 2012.

(3) Summary of Fuel Economy Performance. U.S. Department of Transportation. 2011.

(4) Light-Duty Vehicle Greenhouse Gas Emissions Standards and Corporate Average Fuel Economy Standards; Final Rule. *Federal Register* 75, May 7, 2010, pp 25324−25728.

(5) 2017 and Later Model Year Light-Duty Vehicle Greenhouse Gas Emissions and Corporate Average Fuel Economy Standards; Final Rule. *Federal Register* 77, October 15, 2012, pp 62623−63200.

(6) Gecan, R.; Kile, J.; Beider, P. Effects of Federal Tax Credits for the Purchase of Electric Vehicles. U.S. Congressional Budget Office, 2012; pp 1−36.

(7) Goulder, L.; Jacobsen, M. R.; van Benthem, A. Unintended consequences from nested state and federal regulations: The case of the Pavley greenhouse-gas-per-mile limits. *Journal of Environmental Economics and Management.* **2012**, *63*, 187−207.

(8) Greene, D. Why CAFE Worked. *Energy Policy* **1998**, *26* (8), 595−613.

(9) Azevedo, I. L. Consumer End-Use Energy Efficiency and Rebound Effects. *Annual Reviews of Environment and Resources* **2014**, *39*, 393−418.

(10) Gillingham, K.; Jenn, A.; Azevedo, I. Heterogeneity in the response to gasoline prices: Evidence from Pennsylvania and implications for the rebound effect. *Energy Economics* **2015**, *52*, S41−S52.

(11) Kleit, A. Impacts of Long-Range Increases in the Fuel Economy (CAFE) Standard. *Economic Inquiry* **2004**, *42* (2), 279−294.

(12) Austin, D.; Dinan, T. Clearing the air: The costs and consequences of higher CAFE standards and increased gasoline taxes. *Journal of Environmental Economics and Management* **2005**, *50* (3), 562−582.

(13) Gerard, D.; Lave, L. The economics of CAFE reconsidered: a response to CAFE critics and a case for fuel economy standards. *AEI-Brookings Joint Center for Regulatory Studies, Regulatory Analysis*, 2003; pp 03−01.

(14) Morrow, W. R.; Gallagher, K. S.; Collantes, G.; Lee, H. Analysis of policies to reduce oil consumption and greenhouse-gas emissions from the US transportation sector. *Energy Policy* **2010**, *38* (3), 1305−1320.

(15) Karplus, V.; Paltsev, S. Proposed Vehicle Fuel Economy Standards in the United States for 2017 to 2025. *Transp. Res. Rec.* **2012**, *2287* (1), 132−139.

(16) Cheah, L.; Heywood, J.; Kirchain, R. The energy impact of US passenger vehicle fuel economy standards. 2010 IEEE International Symposium on Sustainable Systems and Technology (ISSST), 2010; pp 1−6.

(17) Bastani, P.; Heywood J., Hope, C. U.S. CAFE Standards, 2012.

(18) Final Rulemaking to Establish Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards Regulatory Impact Analysis. U.S. Environmental Protection Agency, 2010; pp 1.1−9.3.

(19) Ribeiro, K.; Kobayashi, S.; Beuthe, M.; Gasca, J.; Greene, D.; Lee, D. S.; Muromachi, Y.; Newton, P. J.; Plotkin, S.; Sperling, D.; Wit, R.; Zhou, P. J. Transport and its infrastructure. *Climate Change 2007: Mitigation*, 2007.

(20) Whitefoot, K.; Fowlie, M.; Skerlos, S. Product design response to industrial policy: Evaluating fuel economy standards using an engineering model of endogenous product design. Energy Institute at Haas Working Paper WP-214, 2011.

(21) Shiau, C. N.; Michalek, J.; Hendrickson, C. A structural analysis of vehicle design responses to Corporate Average Fuel Economy policy. *Transportation Research Part A: Policy and Practice* **2009**, *43* (9), 814−828.

(22) Whitefoot, K.; Skerlos, S. Design incentives to increase vehicle size created from the US footprint-based fuel economy standards. *Energy Policy* **2012**, *41*, 402−411.

(23) Anderson, S.; Sallee, J. Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards. *American Economic Review* **2011**, *101* (4), 1375−1409.

(24) 2012 Annual Energy Outlook. U.S. Energy Information Administration, 2012.

(25) 2013 Annual Energy Outlook. U.S. Energy Information Administration, 2013.

(26) 2014 Annual Energy Outlook. U.S. Energy Information Administration, 2014.

(27) 2015 Annual Energy Outlook. U.S. Energy Information Administration, 2015.

(28) 2009 National Household Travel Survey. U.S. Department of Transportation, Federal Highway Administration, 2009.

(29) Fuel Economy Datafile. U.S. Department of Energy and U.S. Environmental Protection Agency, 2015. http://www.fueleconomy.gov/feg/download.shtml (accessed December 8, 2015).

(30) Siler-Evans, K.; Azevedo, L.; Morgan, G. Marginal emissions factors for the US electricity system. *Environ. Sci. Technol.* **2012**, *46* (9), 4742−4748.

(31) Siler-Evans, K.; Azevedo, I.; Morgan, G.; Apt, J. Regional variations in the health, environmental, and climate benefits of wind and solar generation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (29), 11768−11773.

(32) Graff Zivin, J.; Kotchen, M.; Mansur, E. Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies. *Journal of Economic Behavior & Organization* **2014**, *107*, 248−268.

(33) Anair, D., Mahmassani, A. State of Charge: Electric Vehicles' Global Warming Emissions and Fuel-Cost Savings across the United States, Report: Union of Concerned Scientists, June 2012. http://www.ucsusa.org/sites/default/files/legacy/assets/documents/clean_vehicles/electric-car-global-warming-emissions-report.pdf (accessed October 8, 2015).

(34) Patil, R.; Adornato, B.; Filipi, Z. Design optimization of a series plug-in hybrid electric vehicle for real-world driving conditions. *SAE International Journal of Engines* **2010**, *3* (1), 655−665.

(35) Tamayao, M.; Michalek, J.J.; Hendrickson, C.; Azevedo, I. Regional variability and uncertainty of electric vehicle life cycle CO2 emissions across the United States. *Environ. Sci. Technol.* **2015**, *49*, 8844−8855.

(36) Michalek, J. J.; Chester, M.; Jaramillo, P.; Samaras, C.; Shiau, C. S.; Lave, L. Valuation of plug-in vehicle life cycle air emissions and oil displacement benefits. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (40), 16554−16558.

(37) Tessum, C.; Hill, J.; Marshall, J. Life cycle air quality impacts of conventional and alternative light-duty transportation in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (52), 18490−18495.

(38) Greene, D.; Park, S.; Liu, C. Analyzing the Transition to Electric Drive in California. *Futures* **2014**, *58*, 34−52.

(39) Transitions to Alternative Vehicles and Fuels. National Research Council of the National Academies, 2013.

(40) Revkin, A. No Red and Blue Divide When it Comes to Renewable Energy Innovation and CO_2 Rules. The New York Times, April 6, 2015. http://dotearth.blogs.nytimes.com/2015/04/06/no-red-and-blue-divide-when-it-comes-to-renewable-energy-innovation-and-co2-rules/.

DOI: 10.1021/acs.est.5b02842
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

The Rebound Effect and Energy Efficiency Policy

Kenneth Gillingham*, David Rapson†, and Gernot Wagner‡

Introduction

Buy a more fuel-efficient car, drive more. This is perhaps the simplest illustration of what has come to be known as the rebound effect—the phenomenon that an increase in energy efficiency may lead to less energy savings than would be expected by simply multiplying the change in energy efficiency by the energy use prior to the change. The existence of the rebound effect has been clear for a long time. In fact, Jevons (1865) hypothesized that greater energy efficiency may even lead to a "backfire," whereby industrial energy use increases. However, the size of the rebound effect is much less clear. There is great variation in estimates, which stems from differences in definitions of the rebound effect, as well as in the quality of the data and the empirical methodologies used to estimate it. This has clear policy implications because both researchers and policymakers need reliable information about the magnitude of the rebound effect to evaluate the energy savings and economic welfare implications of energy efficiency policies. Although the rebound effect is just one component of this more important analysis, it has received significant attention, including in the popular media, which is often in search of counterintuitive results.

The goal of this article is to more clearly define the rebound effect in the context of energy-efficiency improvements, including clarifying its various channels, and to critically assess the literature that estimates its magnitude. In particular, we distinguish between the rebound effect from a costless exogenous energy efficiency improvement—what we will refer to here as a zero-cost breakthrough—and the rebound effect from an actual (typically costly) energy efficiency policy—what we will refer to here as a policy-induced improvement. Recognition of this

* Yale University and the National Bureau of Economic Research; e-mail: kenneth.gillingham@yale.edu.

† University of California, Davis; e-mail: dsrapson@ucdavis.edu.

‡ Environmental Defense Fund and School of International and Public Affairs, Columbia University; e-mail: gernot@gwagner.com.

The research for this article arose from an Environmental Defense Fund workshop. Our thanks go to Matthew Kotchen, who helped conceptualize this project. We thank Blake Alcott, Severin Borenstein, Dallas Burtraw, Nathan Chan, Frank J. Convery, Manuel Frondel, Nathaniel O. Keohane, Reid Leifset, Derek Lemoine, Kenneth Small, Thomas Sterner, Karen Turner, Colin Vance, Matthew Zaragoza-Watkins, Richard J. Zeckhauser, the editor, and two anonymous referees for comments, as well as Jonathan Camuzeaux, Ruiwen Lee and Paige Weber for excellent research assistance. All remaining errors are our own. The authors do not declare any conflict of interest.

Review of Environmental Economics and Policy, volume 10, issue 1, Winter 2016, pp. 68–88
doi:10.1093/reep/rev017
Advance Access published on December 14, 2015

distinction can be helpful for interpreting estimates in the literature, which often conflate the two, leading to inappropriate conclusions and an exaggerated rebound effect.

The most common approach in the literature for estimating the rebound effect is to empirically estimate fuel-price or operating-cost elasticities of demand. However, such estimates should be treated with caution precisely because they conflate the zero-cost breakthrough and policy-induced improvement effects. When we consider the cumulative rebound effect, especially if we include rebound effects that may occur at the macroeconomic level, reliable empirical estimates are much harder to come by.

The article is structured as follows. First we define the different components of the rebound effect. Then we review the quantitative evidence in the literature on the different microeconomic and macroeconomic channels of the rebound effect and discuss challenges to identifying causal rebound effects for each channel. We conclude with a discussion of the implications of the rebound effect for energy efficiency policy.[1]

Defining the Rebound Effect

The classic way that researchers have approached the rebound effect in the literature has been to consider an improvement in energy efficiency and then compare the achieved reductions in energy use to those forecasted without any consumer and market responses to the energy-efficiency improvement. Such consumer and market-wide responses are likely to occur because the energy efficiency improvement itself changes relative prices (and, thus, real income). The rebound effect is then expressed as the percentage of the forecasted reduction in energy use that is lost due to the sum of the consumer and market responses.

To illustrate, consider an air conditioner with annual electricity use of 100 kWh/year. Suppose a more efficient air conditioner shaved 10 kWh/year off this total before accounting for any consumer and market responses. If these responses increased electricity use by 1 kWh/year, then the rebound effect would be equal to 10 percent—that is, 1 of the 10 kWh per year in expected energy savings would be "taken back" due to the consumer and market responses.[2]

Exogenous versus Bundled Improvements in Energy Efficiency

Although this broad definition captures the essence of the rebound effect, it neglects the way in which energy efficiency is actually improved. The literature makes different assumptions about this key issue, which can cause misconceptions about exactly what the rebound effect is, how to estimate it, and how to interpret those estimates. It is helpful to begin with the distinction between (1) an exogenous increase in energy efficiency (holding other product attributes constant) and (2) a change in energy efficiency that is bundled with changes in other product attributes (e.g., a more energy-efficient air conditioner that is also smaller overall and, thus, can

[1]Throughout the article, we highlight common misconceptions about the rebound effect and how to address them. See the online supplementary materials for a table that summarizes the main misconceptions about the rebound effect.

[2]Here we follow the literature by defining the rebound effect with respect to energy. One could analogously define the rebound effect with respect to emissions (Thomas and Azevedo 2013), which in many cases is proportional to the energy rebound. Exceptions include biofuels policies that lead to indirect land use emissions or policies that lead to fuel switching, for example from coal to natural gas and, thus, from carbon to methane emissions.

work in different windows), which may induce a change in the energy service provided and perhaps also in the cost of the product.[3]

To illustrate this distinction, first consider an exogenous increase in energy efficiency—a zero-cost breakthrough—in which an innovation allows a product (e.g., an appliance) manufacturer to increase energy efficiency costlessly while holding all other attributes of the product the same. The resulting consumer and market responses are a pure rebound effect because they capture only those responses induced by the improvement in energy efficiency.

In contrast, consider a policy-induced improvement, whereby a policy requires manufacturers to improve the energy efficiency of a particular product. In this case, the energy efficiency improvement may be costly, potentially raising the price of the product. At the same time, the policy may induce or even necessitate changes in other attributes of the product, such as size, weight, or capacity. In this case, both the price of the product and the energy service it provides may change along with the improvement in energy efficiency.[4]

Thus, for both estimation and policy purposes, it is crucial to distinguish between zero-cost breakthroughs and policy-induced improvements. If we are seeking to estimate a response attributed directly to an energy efficiency improvement, then the zero-cost breakthrough approach is likely to be a better measure of the rebound effect. Any empirical estimation that controls for all of the key attributes of a product is aiming to identify this pure effect. In fact, this is the most common approach used to estimate what most researchers call the rebound effect.

In contrast, if we are interested in the overall effect of a policy—the bundle of changes that occurs, including but not limited to energy efficiency—then focusing on a policy-induced improvement is the appropriate approach. In this case, the goal would be to estimate a compound effect that combines the energy savings from the efficiency improvement with the energy adjustments due to changes in the attributes and cost of the product. This estimate may even capture changes in sales of the product or other consequences. To calculate the policy-induced improvement rebound effect, one could examine the difference in the forecasted energy savings (based on a simple engineering calculation) and the empirically estimated effect. This result may be appropriate for considering the energy implications of a specific policy but is generally not equivalent to the more pure concept of the rebound effect represented by the zero-cost breakthrough approach.

Which Is the Preferred Approach for Policy Analysis?

Neither the zero-cost breakthrough nor the policy-induced improvement approach is unambiguously a better choice for policy analysis. The choice depends on context and the specific question at hand. The zero-cost breakthrough approach, which isolates the effect of an exogenous energy efficiency improvement on the consumer and market responses, provides clear guidance on how changes in energy efficiency alone would change energy use. These results

[3]Energy is a demand that is derived from the consumers' demand for energy services (e.g., miles driven in a particular car, refrigeration). These energy services themselves may change along with the attributes of a product (e.g., a refrigerator with an ice maker provides a different energy service than a refrigerator without an ice maker).

[4]There may be a continuum between zero-cost breakthroughs and policy-induced improvements, whereby a rebound effect captures some, but not all, of the changes from a policy. However, such intermediate cases may be more difficult to interpret in terms of policy implications. Thus, we focus our discussion here on the two extremes: zero-cost breakthrough and policy-induced improvement.

are likely to be more widely applicable than focusing on a specific policy-induced improvement because the approach holds constant potentially confounding variables. Thus, the results can be used to establish the degree to which the rebound effect improves social welfare by providing cheaper energy services that consumers value. Moreover, if policy-induced energy efficiency improvements are associated with only negligible costs and changes in attributes, then estimates for zero-cost breakthrough may be similar to those for policy-induced improvements. However, in most cases, an energy efficiency policy also causes changes in costs and attributes. It is difficult to disentangle these responses empirically because it is essential to know all of the pertinent consumer and market responses to the improved efficiency, the changes in attributes, and the increased cost of the product itself. All of these responses (which comprise the policy's overall effect) play a role in what ultimately matters most to policymakers: the energy efficiency policy's effects on social welfare.

Microeconomic Channels for the Rebound Effect

Before moving to empirical estimates of zero-cost breakthroughs and policy-induced improvements, it is useful to review some basic microeconomic theory to highlight the channels by which the microeconomic rebound occurs. These channels stem from the classic substitution and income effects of consumer theory. We focus only on consumer theory here but address rebound effects from producers in our discussion of the macroeconomic rebound.

Substitution and Income Effects

When energy efficiency improves, the price of energy services changes. Substitution and income effects arise, which influence consumers' consumption of the energy services and, ultimately, energy use. Measuring these effects is not straightforward. In the case of a zero-cost breakthrough, the decline in the cost of the energy services implies that consumers will make a series of four adjustments to their consumption bundle,[5] which may, in turn, affect their derived demand for energy. First, consumers will substitute toward the more energy-efficient product, which is now relatively less expensive. Second, consumers will substitute away from other now relatively more expensive goods.[6] Third, the lower effective price for the energy service increases the consumer's purchasing power, which means consumers will further increase consumption of the more energy-efficient product (assuming it is a normal good). Finally, their increased purchasing power means that consumers will also increase their consumption of other normal goods. Each of these adjustments will either increase or decrease the amount of energy used for the consumer's consumption bundle.

The direct rebound effect

These four effects do not perfectly match the terms most commonly used in the literature on the rebound effect. The direct rebound effect is generally defined as the change in energy use resulting from the combined substitution and income effects on the demand for the energy-efficient product

[5]See Borenstein (2015) for a more technical discussion of these channels of behavioral adjustment.
[6]More broadly, consumers will change their bundle of consumption toward complements to (and away from substitutes for) the energy-efficient product.

(Sorrell and Dimitropoulos 2008). This definition is convenient because economists typically estimate elasticities of demand (e.g., the marginal change in demand for air conditioning as the operating cost of the air conditioner changes), which can be easily converted into a direct rebound effect. Using these elasticity estimates implicitly adopts the zero-cost breakthrough approach to the rebound effect because it tells us, for example, how much additional air conditioning consumers will use if their operating cost changes on the margin, holding all other product attributes constant. For example, if the elasticity of demand with respect to the operating cost is −0.5, then 50 percent of the reduction in energy use from an improvement in energy efficiency on the margin will be taken back by the substitution and income effects, which increases the energy use.[7] It is important to note that this estimate of the direct rebound effect ignores any changes in the demand for other goods due to either the change in relative prices or purchasing power. Nonetheless, the direct rebound effect is useful for quantifying and understanding the first-order consumer response to an increase in energy efficiency.

The indirect rebound effect

The effect of an energy efficiency increase on the demand for all other goods and the subsequent change in energy use is called the indirect rebound effect. However, the literature is not consistent in how this term is used. Some studies include any changes in energy use resulting from changes in the demand for other goods, including substitution effects, income effects, and any embodied energy used to create the energy efficiency improvement (Azevedo 2014). Other studies use the term indirect rebound effect even more broadly, by including substitution effects, income effects, embodied energy, and even macroeconomic rebound effects (Sorrell and Dimitropoulos 2008). However, the most common approach in the literature is to refer to the indirect rebound effect as including only the income effects on the consumption of all other goods. For example, buyers of a more fuel-efficient vehicle may decide to spend the savings on a flight for a vacation—another energy-intensive activity—or on something much less energy intensive, such as books and movies. The sign and magnitude of this indirect rebound effect depends on the difference in energy intensity (per dollar) between the energy-efficient product (prior to the efficiency improvement) and other goods consumed on the margin. It is important to recognize that this more common definition of the indirect rebound effect ignores the substitution effects on other goods that arise from the decrease in the cost of using the more energy-efficient product.[8] Along the same lines, the literature commonly ignores any cost of the efficiency improvement, even though such a cost would produce income effects—reducing (increasing) the indirect rebound effect if, before the improvement, the energy-efficient product is more (less) energy intensive than the marginal consumption bundle (Borenstein 2015).

The microeconomic rebound and welfare

The income and substitution effects described here are no different from any other adjustments that consumers make when confronted with a change in relative prices. By revealed preference, consumers are enjoying private surplus gains. Thus, it follows that a net welfare decrease from a rebound effect is only possible if the external costs associated with these adjustments to the consumer's consumption bundle outweigh the private gains. For example, the external

[7]Note that this approach ignores the substitution and income effects on other goods.
[8]These substitution effects are typically implicitly assumed away as being insignificant.

pollution costs from particularly dirty electricity use could outweigh the consumer surplus benefits from consumers increasing usage of a more efficient air conditioner and reoptimizing their consumption bundle.[9]

Estimating Microeconomic Rebound Effects

We now turn to estimation. Because the microeconomic rebound effect consists of substitution and income effects across all goods, an attempt to fully measure the rebound effect would require estimating the substitution and income effects for all goods in the economy—clearly an infeasible task. Instead, most studies ignore the demand for other goods and focus on estimating the price elasticity of demand for the more energy-efficient product—the zero-cost breakthrough approach. A few studies estimate the effect of a policy—the policy-induced improvements approach—although again they generally ignore effects on other goods in the economy. There are also a few estimates of the income effects from changing the energy consumption of all other goods, but these are generally based on the average rather than the marginal consumption bundle. We are not aware of any studies that estimate these own- and other-good effects jointly using comparable data sources. This may bias rebound effect estimates because a greater increase in demand for the energy-efficient product (i.e., direct rebound) generally implies a smaller increase in demand for other goods (i.e., substitution and income effects on other goods) (Chan and Gillingham 2015).

Caveats

Before discussing specific estimates, additional caveats are in order. First, to provide reliable guidance for analyses, it is critical that studies estimate a causal effect. This is particularly important when using demand elasticities to quantify the rebound effect.[10] For example, studies that rely on cross-sectional variation in fuel prices or operating costs may have difficulty controlling for unobserved heterogeneity. Such studies, even if otherwise well executed, tend to find much more elastic demand than studies that include other sources of variation (e.g., see West 2004).

Second, the conversion of a demand elasticity into an estimate of the direct rebound effect requires an assumption about symmetry of consumer response to changes in fuel prices and energy efficiency. Under standard neoclassical assumptions, the utilization of an energy-consuming good is based on the operating cost (i.e., the fuel price divided by the energy efficiency). Therefore, a change in both the fuel price and in the energy efficiency of the good will change the operating cost in identical (but opposite) ways. Thus, it is common in the literature to describe the fuel-price elasticity of demand as being the direct rebound effect, as we will see. However, in settings where multiple energy services use the same fuel, the fuel-price elasticity and the direct rebound effect are not one and the same (Chan and Gillingham 2015). Furthermore, recent evidence concerning passenger transportation suggests that consumers may respond less to changes in energy efficiency than to changes in fuel price (Gillingham 2011). This may occur because fuel prices are more salient: consumers see them every time they

[9]See Chan and Gillingham (2015) for a detailed examination of these welfare effects.
[10]Many studies estimating demand elasticities do not meet current standards for identification and fail to address standard endogeneity issues such as simultaneity.

pay their energy bill. In this case, using the fuel-price elasticity of demand would overestimate the direct rebound effect. However, other studies show either no asymmetry in response (Frondel and Vance 2013) or a greater response to changes in energy efficiency than to changes in fuel price (Linn 2013). One potential explanation for a greater response to changes in energy efficiency is the perceived longevity of such changes. Li, Linn, and Muehlegger (2014) find that gasoline taxes appear to be more salient than fuel prices, perhaps again due to perceived longevity. Thus, further research is needed into the symmetry of fuel-price elasticities and energy efficiency elasticities.

Third, the consumer response to any change in usage costs may vary depending on the timeframe of the response. For example, when fuel prices change, in the short run consumers can choose how many trips to take, what route to take, which vehicle to take (if they have multiple vehicles), and whether to take public transportation (if available). In the medium run, they can purchase or scrap vehicles, and in the long run they can choose where to live and work. It is likely that long-run energy demand is more elastic than short-run demand; yet long-run elasticities are harder to estimate credibly and thus harder to come by.

Finally, each estimate of price elasticities is for a particular time and place, and energy demand could vary with the specific setting. For example, Gillingham (2014) shows that the elasticity of demand for driving with respect to the price of gasoline exhibits noticeable heterogeneity across different counties in California. One could imagine that there would be even greater differences when examining a developing country or a country with an extensive public transportation system. The bottom line here is that even if an elasticity estimate is internally valid, we need to examine its external validity before applying it elsewhere.

With these caveats in mind, we next review the relevant elasticity estimates in the literature that may be useful in providing policy guidance to economists and policymakers.

Elasticities for Developed Countries

We first discuss the literature for developed countries. Given the vast number of estimates, we present selected reliable estimates, with a focus on studies of overall demand or household-level demand (table 1).[11]

The studies we include in table 1 were selected because they are more recent and use rigorous empirical methods such as panel data methods, experimental designs, or quasi-experimental designs. These studies attempt to address potential endogeneity concerns and present some evidence of internal validity. They tend not to rely exclusively on cross-sectional variation. All provide either short-run or medium-run estimates. As emphasized by Hamilton (2009) and Gillingham (2011), including a lagged dependent variable to distinguish between short-run and long-run responses requires strong assumptions. Yet, nearly all estimates of long-run responses are based on either an ordinary least squares regression with a lagged dependent variable or on cross-sectional variation (with the assumption that it is capturing a long-run equilibrium). Thus, we believe that the short-run and medium-run estimates are more reliable.

The primary theme that emerges from our review of this literature is that the short-run and medium-run elasticities of demand for gasoline/driving and electricity are generally in the range

[11]For more comprehensive reviews of estimates of elasticities in different sectors, see Greening, Greene, and Difiglio (2000), Sorrell (2007), Jenkins, Nordhaus, and Shellenberger (2011), and Gillingham (2011). Not surprisingly, these reviews show large ranges of estimates in most sectors.

Table 1 Selected elasticity estimates for developed countries

Study	Type of price elasticity	Estimated value
Allcott (2011)	Illinois short-run elasticity of electricity demand, 2003 and 2004	−0.1
Barla et al. (2009)	Canada short-run elasticity of VMT demand, 1990–2004	−0.08
Frondel and Vance (2013)	Germany short-run elasticity of VMT demand, 1997–2009	−0.458[a]
Gillingham (2014)	California medium-run new vehicle elasticity of VMT demand, 2001–2009	−0.23
Hughes, Knittel, and Sperling (2008)	U.S. short-run elasticity of gasoline demand, 1975–1980	−0.21 to −0.34
Hughes, Knittel, and Sperling (2008)	U.S. short-run elasticity of gasoline demand, 2001–2006	−0.034 to −0.077
Ito (2014)	California medium-run elasticity of electricity demand, 1999–2007	−0.088
Jessoe and Rapson (2014)	Connecticut short-run elasticity of electricity demand, 2011	−0.12
Small and van Dender (2007)	U.S. short-run elasticity of VMT demand, 1966–2001	−0.045[b]

Notes: All electricity demand elasticity estimates are for residential customers. VMT refers to vehicle miles traveled.
[a]We report the fixed effects estimate, which we believe to be the most reliable.
[b]We use the estimate from the 1997–2001 period; earlier elasticities were higher in absolute value.

of −0.05 to −0.40, suggesting a direct rebound effect on the order of 5 percent to 40 percent, with most of the studies falling in the range of 5 percent to 25 percent. All of these studies focus on gasoline or electricity use, and it may not be appropriate to apply the estimates to other energy services, including those that use natural gas, heating oil, or other fuels. Unfortunately, there is scant evidence on the price elasticity of demand for other energy services; all of the published papers we could find are more than a decade old and use limited data. In a review of the older literature, Sorrell (2007) finds wide ranges for most residential energy services. Thus, we believe that new research is needed on these other energy services. Moreover, new studies are needed to help us identify the size of the error from using own-price elasticities for the direct rebound.

Most of the studies cited in table 1 are for the United States. Because each country has unique circumstances, it may be inappropriate to apply the estimates in table 1 to other regions and countries, both developed and developing.[12]

Elasticities for Developing Countries

For developing countries, one might hypothesize a greater elasticity of demand to price changes, and thus direct rebound effect, because of the greater unmet demand for energy services. However, there are a variety of country-specific factors that may affect responsiveness in any given market, such as the wealth of those who own vehicles or appliances. In our review

[12]For example, Frondel et al. (2013), which uses data for Germany—a country with better public transportation and higher gasoline prices than the United States—finds a more elastic response in driving to changes in gasoline prices than the other studies in table 1.

of the literature, we found a surprising number of studies estimating elasticities of usage for durable goods in low- and middle-income countries. However, the authors of these studies often face severe data limitations and measurement error in the data. Moreover, these studies rarely meet current standards for identification in applied economics, and the caveats above certainly apply here.

Table 2 shows a representative sample of studies published in peer-reviewed journals. We have not screened these studies for reliability (as we did for the developed countries) because nearly all of them face data limitations. We should, thus, be very cautious in viewing them as causal estimates of price elasticities. These estimates of demand elasticities in developing countries range widely, with the most common range on the order of −0.10 to −0.40 in the short run. Despite the limitations of some of these studies, it is interesting to note that the estimated

Table 2 Representative sample of recent price elasticity estimates for low- and middle-income countries

Study	Type of elasticity	Estimated value
Al-Faris (2002)	Gulf Cooperation Council short-run elasticity of total electricity demand, 1970–1997	−0.09
Alves and De Losso da Silveira Bueno (2003)	Brazil short-run elasticity of gasoline demand, 1974–1999	−0.09
Atakhanova and Howie (2007)	Kazakhstan short-run elasticity of electricity demand, 1994–2003	−0.128[a]
Athukorala and Wilson (2010)	Sri Lanka short-run elasticity of total elasticity demand, 1960–2007	−0.16
Ben Sita, Marrouch, and Abosedra (2012)	Lebanon short-run elasticity of gasoline demand, 2000–2010	−0.623
Crotte, Noland, and Graham (2010)	Mexico short-run elasticity of gasoline demand, 1980–2006	0 to −0.15
Halicioglu (2007)	Turkey short-run elasticity of electricity demand, 1968–2005	−0.33 to −0.46
Iwayemi, Adenikinju, and Babatunde (2010)	Nigeria short-run elasticity of gasoline demand, 1976–2006	−0.25
Jamil and Ahmad (2011)	Pakistan short-run elasticity of total electricity demand, 2000s	−0.07
Lin and Zeng (2013)	China medium-run elasticity of gasoline demand, 1997–2008	−0.196 to −0.497
Nahata et al. (2007)	Russia short-run elasticity of electricity demand, 1995–2000	−0.165 to −0.28
Ramanathan (1999)	India short-run elasticity of gasoline demand, 1972–1993	−0.21
Sene (2012)	Senegal short-run elasticity of gasoline demand, 1970–2008	−0.12
Zein-Elabdin (1997)	Sudan short-run elasticity of charcoal demand, 1960–1990	−0.55
Ziramba (2008)	South Africa short-run elasticity of electricity demand, 1978–2005	−0.02

Notes: Gulf Cooperation Council countries are Saudi Arabia, Kuwait, Bahrain, Qatar, United Arab Emirates, and Oman. All electricity demand elasticity estimates are for residential customers unless otherwise noted.
[a]We report the instrumental variables fixed effects estimate.

elasticities for developing countries are in the same range as the estimates for developed countries.

Estimated Policy-Induced Improvements

Estimating the rebound effect for policy-induced improvements requires more than just the fuel-price elasticity of demand because other product attributes may also have changed. Recent studies have used variation from natural experiments to estimate rebound effects in this context.

For example, in a field experiment in which households are given more efficient clothes washers, Davis (2008) finds a price elasticity of clothes washing of −0.06. This estimate is similar to a zero-cost breakthrough but with a key difference: the new clothes washers given to the households were larger and gentler on clothes than the old washers. This means that households may have adjusted their clothes washing behavior in response not only to the change in the price of the energy service but also to improved nonprice product attributes. In fact, the increase in clothes washer use resulted from households running more clothes in each wash. This estimate is capturing the direct rebound effect of a policy-induced improvement. That is, it captures the effects from both the change in energy efficiency and the change in the quality of the energy service (i.e., clothes washing).

In a similar study, Davis, Fuchs, and Gertler (2015), examine a program in Mexico that provides direct cash payments and subsidized financing to consumers replacing old air conditioners and refrigerators with new energy-efficient appliances, much like the cash-for-clunkers program for vehicles in the United States. They find that electricity use dropped by only 7 percent after replacing the old refrigerator with a new, efficient one and that electricity use actually increased after replacing an air conditioner. These results suggest a potentially very large change in the energy service (e.g., the new refrigerators may have been much larger or the air conditioners quieter), as well as an income effect from the transfer, which together lead to a large apparent rebound effect from this policy.

Finally, Gillingham (2013) examines the direct rebound effect of a policy-induced change in vehicle prices that leads to consumers purchasing different vehicles (each with bundles of attributes) and then driving them more. The result is an elasticity of driving with respect to operating costs of −0.15 for new vehicles in California. We believe that further research on the rebound effect of policy-induced improvements is very important for policy.

Estimates of Rebound Effects on Other Goods

As mentioned earlier, changing the energy efficiency of a good may affect overall energy demand through changes in the demand for other goods in the consumption bundle, which occur through the substitution and income effects on these goods. Most studies seek to estimate only the income effects for other goods (calling this the indirect rebound) by answering the question: If consumers are given an extra dollar, how will they spend it?[13] One approach has been to assume that consumers make purchases associated with the average energy intensity of all consumer goods, which is often referred to as proportional respending. Studies that follow

[13]Specifically, we would want to know how consumers would spend the dollar on all goods except the more energy-efficient one.

this approach generally examine the energy intensity of the economy using either input-output tables or other aggregate statistics of economic activity and energy use. A second approach is to use cross-sectional data to compare consumption patterns across income brackets (Thiesen et al. 2008). A third approach is to use income elasticities that are based on how consumers' demand for goods changes over time as income rises (Druckman et al. 2011). The findings in this literature vary, but most recent studies tend to estimate a consumption elasticity with respect to income on the order of 5 percent to 15 percent (Druckman et al. 2011; Thomas and Azevedo 2013). Thomas and Azevedo (2013) also make assumptions in order to bound the estimated substitution effects for other goods. One would expect that these effects would vary depending on the cross-elasticities between the good in question and other energy-using goods, the additional cost of the more efficient good, and any additional energy use from the production of the more efficient good. It is important to note that all existing estimates assume a zero-cost breakthrough scenario. Any additional costs would reduce the income effects on other goods, thus reducing the indirect rebound. In addition, most existing estimates are for developed countries, although there has been some work on the income elasticity of energy use in developing countries (see, e.g., Wolfram, Shelef, and Gertler 2012).

Macroeconomic Channels for the Rebound Effect

The macroeconomic rebound effect is complex. This is because markets re-equilibrate when the demand for an energy resource changes, and an increase in energy efficiency may affect overall energy demand through several channels of adjustment. In this section we seek to clarify this issue in four ways: (1) we define the macroeconomic rebound and review the theoretical pathways that are thought to generate it; (2) we describe the challenges inherent in trying to quantify the magnitude of the macroeconomic rebound, including discussing common pitfalls; (3) we review what the theoretical and empirical literature tells us about the potential magnitude of the macroeconomic rebound; and (4) we discuss what this means for environmental economics research and policymaking.

Defining Macroeconomic Rebound Effects

The literature defines the macroeconomic rebound effect as an increase in energy use after an energy efficiency improvement through market adjustments and innovation channels. Such an effect is easiest to consider in the context of a zero-cost breakthrough, which underpins much of the discussion that follows.[14] We divide our discussion into a macroeconomic price effect and a macroeconomic growth effect.

Macroeconomic price effect

The macroeconomic price effect is an economy-wide analog to the microeconomic direct rebound effect that works through prices (Gillingham et al. 2013). When an energy efficiency improvement shifts the market demand curve for energy down (i.e., to the left), consumers and producers will adjust until a new equilibrium is reached. To illustrate, consider the global oil

[14]Although in theory it is possible to consider macroeconomic rebound effects in the context of a policy-induced improvement, we have never seen this done in practice.

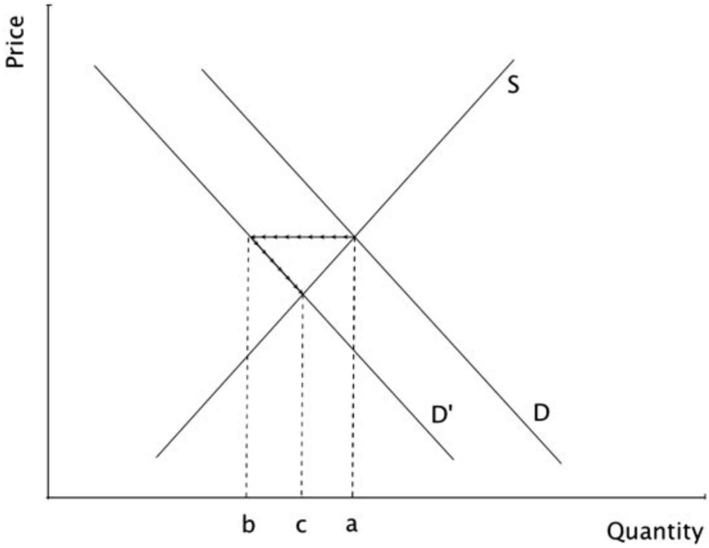


Figure 1 Macroeconomic price effect
Notes: This graph depicts the macroeconomic price effect associated with an energy efficiency policy that shifts demand inwards from D to D′. The shift in demand reduces the quantity of energy demanded from a to b; however, the equilibrium outcome yields a smaller energy reduction due to the price effect, which moves quantity from b to c.

market. An efficiency improvement in, say, the United States, will lower the global oil price, which increases the global quantity of oil demanded. As shown in figure 1, the initial increase in energy efficiency shifts the global demand curve down, from *D* to *D'*. Because *a* minus *b* is the shift in demand and *a* minus *c* is the change in equilibrium quantity, the macroeconomic price effect is 1–(*a*-*c*)/(*a*-*b*). The magnitude of this rebound effect is thus a function of the slopes of the demand and supply curves, whereby increasingly inelastic supply and increasingly elastic demand induce a higher rebound.

Macroeconomic growth effect

The macroeconomic growth effect, which is often cited but poorly defined, is the rationale behind many of the backfire claims in the literature—that is, that energy efficiency improvements will actually increase energy use.[15] In fact, the classic example given by Jevons (1865) postulates a type of macroeconomic growth effect. The basic premise is that an increase in the efficiency of energy-consuming durables may spur economic growth—and that economic growth requires additional energy consumption. There are three main channels through which a change in energy efficiency could lead to the macroeconomic growth effect.

First, sectoral reallocation may occur due to a change in the relative returns of economic sectors.[16] For example, a change in the productivity of energy inputs in an energy-intensive

[15]For example, Jenkins, Nordhaus, and Shellenberger (2011) state: "The more efficient production and use of energy at a macroeconomic scale drives economic productivity overall and encourages the substitution of energy for other factors of production (e.g., labor), resulting in more rapid economic growth and energy consumption ('macroeconomic rebound' effects)."
[16]Sectoral reallocation in response to changing costs is analogous to a reallocation of inputs into aggregate production in response to changing costs.

sector may improve the relative return on investment in that sector, leading that sector to grow relative to others. This can be (roughly) thought of as the supply-side analogy to the substitution effects discussed in the context of the microeconomic rebound.

A second potential channel is induced innovation—that is, a shock to total factor productivity. One possibility is that an energy efficiency policy (a policy-induced improvement) leads manufacturers to update their processes, thus inducing innovation. Alternatively, a zero-cost breakthrough in one sector may spill over to others. For example, the development of lighter-weight aircraft to improve aircraft efficiency may spill over to other sectors and lead to lighter-weight vehicles. Of course, to be considered a rebound, the innovation in other sectors must be directly attributable to the spillovers from the energy efficiency improvement. Should such spillovers exist, they could increase or decrease energy use in the other sectors.

The third potential channel for the macroeconomic growth effect concerns the deployment of inframarginal resources (i.e., money in the economy that would previously have been spent on energy) that are freed by a zero-cost breakthrough. These may be subject to a fiscal multiplier (see, e.g., Ramey 2011). That is, dollars that were previously spent on energy can now be spent in ways that engage new economic activity that utilizes previously idle resources. Surplus created from this new activity may cause the overall economic impact to exceed the initial amount by some multiplier (Borenstein 2015). Of course, for such a multiplier effect to occur, idle resources must be available so that the incremental resources do not simply crowd out private investment. Although this may be the case during recessions, it is less likely to be the case during economic upswings. More generally, there is strong disagreement among macroeconomists about the size of the fiscal multiplier (Ramey 2011). However, the multiplier in the rebound setting is slightly different because there is long-term debt associated with fiscal stimulus, but not with a zero-cost breakthrough. We are not aware of any study focusing directly on estimating such multipliers in the context of energy efficiency. We turn next to the challenges of estimating macroeconomic rebound effects.

Challenges of Estimating the Macroeconomic Price Effect

The magnitude of the macroeconomic price effect depends on the relative supply and demand elasticities.[17] If the demand elasticity is low and the supply elasticity is high, then the effect will be small. The estimates discussed earlier concerning the price elasticity of gasoline use suggest a relatively inelastic oil demand function, at least in the medium run. The supply of oil is considered to be relatively inelastic in the short run due to capacity constraints. However, oil supply would be expected to be more elastic in the long run because it depends on how development of new extraction technologies responds to price. Unfortunately, there is very little empirical evidence on such supply elasticities. Borenstein (2015) uses oil supply elasticities of 0.2, 0.6, and 1.0 for a sensitivity analysis of the macroeconomic price effect but asserts that the long-run oil supply elasticity may be rather high.[18]

The estimates in Borenstein (2015) indicate that with an oil demand elasticity of −0.4 and an oil supply elasticity of 1.0, the macroeconomic price effect is approximately 30 percent. Using linear demand and supply functions, we arrive at a similar result. However, the possible range

[17]This should be clear from figure 1.
[18]Given the remarkable innovations in oil extraction over the past several decades due to high oil prices, we agree with Borenstein's assertion.

for the macroeconomic price effect is quite large: with a supply elasticity of only 0.2 and demand elasticity of −0.6, we can expect to see a macroeconomic price effect as large as 76 percent. We believe that it is far more likely that long-run oil supply is highly elastic, so we would not expect an effect this large, even if it is possible. Given the likely high long-run oil supply elasticity and low or moderate demand elasticity, we suspect that the macroeconomic price rebound in oil markets is on the order of 20 percent to 30 percent. However, we have not yet seen evidence for other energy markets (e.g., electricity, natural gas). Moreover, for all markets, it is important to recognize that the macroeconomic price effect will always be less than one (demand curves slope downward and supply curves slope upward, by construction). This means that it is theoretically impossible for backfire to occur due solely to the macroeconomic price effect.

Challenges of Estimating the Macroeconomic Growth Effect

Despite being central to backfire claims, the macroeconomic growth effect is the rebound effect topic with the least amount of concrete evidence. Attempts to quantify the macroeconomic growth effect are plagued by the same challenges that are encountered in most macroeconomics research. That is, the global economy is a single, interconnected, complex dynamic system, making definitive arguments about cause and effect nearly impossible. This means, for example, that we cannot say with empirical certainty how U.S. fuel economy standards affect long-run energy use in the United States, let alone in China.

Fortunately, basic economic theory provides some clear guidance on the macroeconomic growth rebound most commonly discussed: sectoral reallocation. The key theoretical insight is that the extent to which a zero-cost breakthrough leads to increases or decreases in overall energy use depends on the elasticities of substitution in consumption and production. To illustrate, consider a household that consumes two goods—an aggregate consumption good (e.g., food or clothing) and an energy service (e.g., driving). This means that households can use their income to purchase either the consumption good or a car and the energy to power it. The question of interest here is: What happens to aggregate energy use in the economy if cars are made more energy efficient?

In the consumer sector, the answer depends on the elasticity of substitution between goods and energy services in the household utility function. To illustrate, let's consider the extremes. If goods and energy services are perfect substitutes, then the household will spend its entire budget on whichever good has the highest utility per dollar spent. If energy services become less expensive than goods (in utility per dollar), then the household may shift its entire budget toward energy services. On the other hand, if goods and energy services are perfect complements, then they will be optimally consumed in fixed proportion. In this case, making one of the goods marginally cheaper (e.g., through energy efficiency standards) will make little difference in consumption and overall energy use because energy is a derived demand (i.e., from energy services). This means that although zero-cost breakthroughs may cause the level of energy services to increase, less energy will be used than before the zero-cost breakthrough. Based on these two extremes, it is clear that there must be a high degree of substitution toward energy services in consumption for the level of actual energy use to increase above pre-energy efficiency improvement levels.

So far, the logic we have presented is the same as the logic behind the microeconomic substitution effects. This means that the consumer substitution effects will be contained in estimates of the sectoral reallocation effect. But sectoral reallocation is even broader; it depends not only on patterns of consumption but also on patterns of production. For production, precisely the same logic applies as for consumption. Where production occurs by combining energy inputs with nonenergy inputs (e.g., capital and labor), the degree of substitutability/complementarity in production determines the overall effect of a zero-cost breakthrough on energy use. If the inputs are highly substitutable, an increase in energy efficiency in production will cause a large swing toward increasing energy inputs. If they are complements, they must be used in fixed proportion, and energy demand will remain unchanged.

A useful implication of these theoretical insights is that the sectoral reallocation rebound is largely driven by the magnitude of substitution elasticities. Intuitively, we would view energy and nonenergy inputs as being more complementary than substitutable in both consumption and production because energy cannot be directly consumed; rather, we use it to help us meet our broader consumption needs. This intuition is shared by Goulder et al. (1999), whose simulation model of alternative abatement policies assumes complementarity of energy and other inputs to production.[19] This leads us to believe that macroeconomic growth rebound effects are likely to be small. However, there is clearly a need for more research to quantify the relevant substitution elasticities.

Empirical Evidence on the Macroeconomic Growth Effect

The theoretical insights just discussed are particularly useful when interpreting the empirical literature on the macroeconomic growth effect, which focuses primarily (but not exclusively) on sectoral reallocation. Other channels may be implicitly included in the macroeconomic growth effect but, to the best of our knowledge, have not been identified separately. There are three strands in the literature that quantify the macroeconomic growth rebound. The first strand uses a structural model of the production function of the economy to make theoretical predictions about the rebound effect. The second attempts to econometrically estimate the total rebound effect (macroeconomic and microeconomic) using historical time-series data. The third involves simulation models of the economy based on input-output tables of economic activity and calibrated relationships between key variables governing economic growth.

Structural models

Beginning with Saunders (1992), there has been a stream of studies in the energy economics literature that relies on a neoclassical growth model to provide theoretical insight into the sectoral reallocation rebound. For example, using a single-sector neoclassical growth model that includes capital, labor, and energy inputs, Saunders (1992) examines how energy efficiency improvements affect overall energy consumption. In this simple setting, the consumer considers energy-intensive goods as perfect substitutes for non-energy-intensive goods. Thus, by construction, Saunders finds that backfire can occur.

Our concern with this and many other models in this literature is that they rely heavily on structural assumptions. For example, switching to a production function that assumes perfect

[19]Goulder et al. (1999) assume an elasticity of substitution of 0.8. It may be even lower in the context here because the energy efficiency intervention itself will already dictate substitution toward more energy-efficient production technology.

complementarity of inputs (i.e., a Leontief production function) would immediately imply zero rebound. Of course, the structural assumption here is just as restrictive as in the single-sector neoclassical growth model. Although such theoretical exercises are interesting, their limitation is that nearly any outcome is possible depending on the choice of structural assumptions and functional forms.

Econometric estimates

Although this observation should not be surprising to macroeconomists, these limitations of structural models have made the use of empirical analyses all the more important for providing reliable guidance on the magnitude of the macroeconomic growth rebound. However, this is where demonstrating causality is critical—but also extremely difficult. For the last century, we have seen large increases in both energy use and the energy efficiency of many durable goods. But in order to claim a causal relationship between energy efficiency and energy use, it must be shown that energy consumption has not increased due to some other factor. Ideally, the experiment needed to identify a zero-cost breakthrough would consist of two worlds—one with the zero-cost breakthrough and one without. Unfortunately, as for many issues in macroeconomics, such an experiment is impossible. In fact, it is extremely difficult, if not impossible, to separate the effect of energy efficiency improvements from exogenous economic growth and the simultaneous dramatic improvements in energy services. Not surprisingly, the few econometric investigations that have relied on historical data to provide evidence of a combined macroeconomic and microeconomic rebound effect leading to backfire (e.g., Tsao et al. 2010; Saunders 2013) have not been published in economics journals, where the standard for empirically identifying a causal effect tends to be higher.

Simulation models

In the absence of credible empirical strategies, macroeconomists often build models of the economy that simulate the effects of policies. This brings us to the third class of approaches used to estimate the macroeconomic rebound effect: calibrated simulation models. These models tend to be general equilibrium models based on input-output tables of economic activity or estimated macroeconometric models with hundreds of equations. Of course, the results of such models are driven by the structure of the model and the parameterization of the relationships. For this reason, many macroeconomic modelers focus on modeling to build intuition, rather than numerical estimates.

The simulation models that are used to numerically estimate the macroeconomic rebound effect compare total energy consumption in a scenario that slightly perturbs the energy efficiency parameter to total energy consumption in the business-as-usual case. If the change in predicted energy use is less than the expected effect of energy efficiency, then the difference is attributed to the rebound effect; if total energy increases, it is consistent with backfire. Some of the most interesting studies in this literature build computable general equilibrium or econometric simulation models of the U.K. economy (e.g., Barker, Ekins, and Foxon 2007; Barker, Dagoumas, and Rubin 2009; Turner 2009). These find results ranging from negative rebounds to massive backfire. This large range of results is very useful for considering the implications of different combinations of structural assumptions and parameter values for the macroeconomic rebound effect. But the reliance of these studies on correlations to parameterize key relationships in the models leaves us unconvinced that they truly pin down the magnitude of the rebound effect. Thus, another valuable area for future research would be analyses that combine clever new empirical approaches with careful numerical simulations.

Implications for Environmental Economics Research and Policy

What does this discussion of the challenges of quantifying the macroeconomic rebound effect tell us about its likely magnitude? Note first that estimates of the sectoral reallocation macroeconomic rebound are not necessarily additive with respect to the microeconomic rebound effects, which are typically already aggregated into the macroeconomic measure. In addition, the macroeconomic price and sectoral reallocation effects may be partly offsetting because sufficiently lower equilibrium energy prices can lead to a reallocation away from energy (Turner 2009). Moreover, to the extent that numerical simulations are based on historical correlations, rather than causal effects, we need to be cautious about interpreting point estimates too literally.

That said, it is possible that there is a substantial macroeconomic growth effect in certain circumstances. Moreover, it appears likely that there is at least some increase in energy consumption from the macroeconomic growth effect, given that it has a theoretically sound basis. Thus, when considering a zero-cost breakthrough, we would recommend that the best current approach for a policy economist would be to calculate the macroeconomic price effect based on the best estimates of elasticities and then perform a sensitivity analysis using different values of the macroeconomic growth rebound effect. Two recent estimates of the macroeconomic growth rebound that could be considered for such a sensitivity analysis are 11 percent (Barker, Ekins, and Foxon 2007) and 21 percent (Barker, Dagoumas, and Rubin 2009).[20] We do not believe that the literature currently provides convincing evidence of a backfire due to the macroeconomic rebound effect.

What does a macroeconomic rebound mean for the welfare effects of policy? The macroeconomic price effect of an energy efficiency improvement arises from reaching equilibria in markets, which improves welfare. Sectoral reallocation leads to more efficient production in an economy, improving welfare. If the energy efficiency improvement induces innovation, this would also improve welfare. However, because these welfare gains may be countered by losses from greater external costs of production or consumption, the net welfare effects are ambiguous.

Conclusions and Implications for Policy

The debate about the magnitude of the rebound effect continues and has important implications for energy efficiency policy. This article has attempted to inform this debate through three main contributions. First, we have introduced the important conceptual distinction between a rebound effect associated with a costless energy efficiency improvement that holds other attributes constant (zero-cost breakthrough) and an energy efficiency policy that may be bundled with other product changes that affect energy use (policy-induced improvement). Second, we have distilled the empirical literature on the microeconomic rebound into a manageable number of estimates that we believe are the most reliable. Third, we have attempted to clarify

[20]This 21 percent is based on the 2020 estimate, whereas the estimate for 2030 is 41 percent. However, both estimates include the income effect within the macroeconomic rebound.

the nature of the macroeconomic rebound and have presented an approach for conceptualizing (or estimating) the size of the effect.

We find that the existing literature does not support claims that energy efficiency gains will be reversed by the rebound effect. Thus we would argue that the continued focus on backfire in policy debates is largely unwarranted and is perhaps distracting attention from the most important issues, such as the welfare implications of energy efficiency policies. In most cases, the total microeconomic rebound has been found to be on the order of 20 percent to 40 percent when all substitution and income effects are included (and perhaps even when the embodied energy in the energy efficiency improvement is included). Far less is known (or knowable) about the macroeconomic rebound. However, we have presented a framework that suggests three conclusions about the macroeconomic rebound. First, although in some markets the macroeconomic price effect may be substantial, it must always be less than 100 percent. Second, the rebound based on sectoral reallocation is likely smaller than the price effect because energy is more likely to be a complement to, rather than substitute for, other inputs in production. Finally, little is known about the effects of induced innovation and productivity on the rebound effect, beyond observing that such developments would almost certainly be welfare increasing. In particular, there is a lack of consensus in the literature that examines how regulation affects total factor productivity. Nevertheless, if induced innovation and productivity lead to a rebound, then quantifying the effect would face the difficult challenge of determining a counterfactual path of innovation and productivity. There is currently scant evidence on this induced innovation channel and thus further research is needed on this topic.

The cumulative effect of these channels of rebound in a zero-cost breakthrough setting may be large in some situations and smaller in others. If pressed to offer our subjective assessment, in most cases we do not expect the total rebound effect to exceed 60 percent, but we recognize that it is possible to have a larger total effect.[21] One might expect a policy-induced improvement to have a larger rebound due to associated changes in product attributes that consumers value, but a smaller rebound to the extent that the cost of the policy mitigates both the income and macroeconomic growth effects. In fact, sufficiently costly energy efficiency policies may well engender negative rebound effects. In sum, while the energy savings from energy efficiency policies will be reduced by the presence of a rebound effect, a zero-cost breakthrough rebound is likely to both conserve energy and increase welfare. The same may be true for a policy-induced improvement rebound, but each policy will require its own analysis.

A primary conclusion of our review is that unless the rebound effect has severe external costs, it will be a benefit, rather than a cost, of an energy efficiency policy. Unfortunately, the focus on minimizing energy use, rather than the broader objective of maximizing economic efficiency, has caused some policymakers to make the mistake of designing policies to mitigate the rebound effect. Such efforts, as discussed in the literature (e.g., van den Bergh 2011) and the policy community (e.g., Gloger 2011), are likely counterproductive from a welfare perspective. Rather than considering the rebound effect as a deterrent to passing energy efficiency policies, policymakers should include the welfare gains and losses as part of their analysis of the benefits of a policy.

[21]This 60 percent estimate is based on a 30 percent long-run microeconomic rebound, 25 percent macroeconomic price effect, and 5 percent macroeconomic growth effect (accounting for the fact that estimates of the macroeconomic growth effect both range widely and may be implicitly including some of the other rebounds).

References

Al-Faris, Abdul Razak. 2002. The demand for electricity in the GCC countries. *Energy Policy* 30 (2): 117–24.

Allcott, Hunt. 2011. Rethinking real-time electricity pricing. *Resource and Energy Economics* 33 (4): 820–42.

Alves, Denisard, and Rodrigo De Losso da Silveira Bueno. 2003. Short-run, long-run and cross elasticities of gasoline demand in Brazil. *Energy Economics* 25 (2): 191–99.

Atakhanova, Zauresh, and Peter Howie. 2007. Electricity demand in Kazakhstan. *Energy Policy* 35 (7): 3729–43.

Athukorala, Wasantha, and Clevo Wilson. 2010. Estimating short and long-term residential demand for electricity: New evidence from Sri Lanka. *Energy Economics* 32 (S1): S34–S40.

Azevedo, Ines L. 2014. Consumer end-use energy efficiency and rebound effects. *Annual Review of Environment and Resources* 39: 393–418.

Barker, Terry, Athanasios Dagoumas, and Jonathan Rubin. 2009. The macroeconomic rebound effect and the world economy. *Energy Efficiency* 2 (4): 411–27.

Barker, Terry, Paul Ekins, and Timothy Foxon. 2007. The macro-economic rebound effect and the UK economy. *Energy Policy* 35 (10): 4935–46.

Barla, Philippe, Bernard Lamonde, Luis F. Miranda-Moreno, and Nathalie Boucher. 2009. Traveled distance, stock and fuel efficiency of private vehicles in Canada: Price elasticities and rebound effect. *Transportation* 36 (4): 389–402.

Ben Sita, Bernard, Walid Marrouch, and Salah Abosedra. 2012. Short-run price and income elasticity of gasoline demand: Evidence from Lebanon. *Energy Policy* 46: 109–15.

Borenstein, Severin. 2015. A microeconomic framework for evaluating energy efficiency rebound and some implications. *Energy Journal* 36 (1): 1–21.

Chan, Nathan, and Kenneth Gillingham. 2015. The microeconomic theory of the rebound effect and its welfare implications. *Journal of the Association of Environmental and Resource Economists* 2 (1): 133–59.

Crotte, Amado, Robert Noland, and Daniel Graham. 2010. An analysis of gasoline demand elasticities at the national and local levels in Mexico. *Energy Policy* 38 (8): 4445–56.

Davis, Lucas. 2008. Durable goods and residential demand for energy and water: Evidence from a field trial. *RAND Journal of Economics* 39 (2): 530–46.

Davis, Lucas, Alan Fuchs, and Paul Gertler. 2015. Cash for coolers: Evaluating a large-scale appliance replacement program in Mexico. *American Economic Journal: Economic Policy* forthcoming.

Druckman, Angela, Mona Chitnis, Steve Sorrell, and Tim Jackson. 2011. Missing carbon reductions? Exploring rebound and backfire effects in UK households. *Energy Policy* 39 (6): 3575–81

Frondel, Manuel, and Colin Vance. 2013. Re-identifying the rebound: What about asymmetry? *Energy Journal* 34 (4): 43–54.

Gillingham, Kenneth. 2011. *The consumer response to gasoline price changes: Empirical evidence and policy implications*. PhD diss., Stanford University.

———. 2013. Selection on anticipated driving and the consumer response to changing gasoline prices. Working paper, Yale University.

———. 2014. Identifying the elasticity of driving: Evidence from a gasoline price shock. *Regional Science & Urban Economics* 47 (4): 13–24.

Gillingham, Kenneth, Matthew Kotchen, David Rapson, and Gernot Wagner. 2013. The rebound effect is over-played. *Nature* 493: 475–76.

Gloger, Stefan. 2011. *Policies to overcome the rebound effect—A new challenge for environmental policy*. Technical Report, Ministry of the Environment, Baden-Wuerttemberg.

Goulder, Lawrence, Ian Parry, Roberton Williams, and Dallas Burtraw. 1999. The cost-effectiveness of alternative instruments for environmental protection in a second-best setting. *Journal of Public Economics* 72 (3): 329–60.

Greening, Lorna A., David L. Greene, and Carmen Difiglio. 2000. Energy efficiency and consumption—The rebound effect—A survey. *Energy Policy* 28 (6–7): 389–401.

Halicioglu, Ferda. 2007. Residential electricity demand dynamics in Turkey. *Energy Economics* 29 (2): 199–210.

Hamilton, James. 2009. Understanding crude oil prices. *Energy Journal* 30 (2): 179–206.

Hughes, Jonathan, Christopher Knittel, and Daniel Sperling. 2008. Evidence of a shift in the short-run price elasticity of gasoline demand. *Energy Journal* 29 (1): 93–114.

Ito, Koichiro. 2014. Do Consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review* 104 (2): 537–63.

Iwayemi, Akin, Adeola Adenikinju, and Adetunji Babatunde. 2010. Estimating petroleum products demand elasticities in Nigeria: A multivariate cointegration approach. *Energy Economics* 32 (1): 73–85.

Jamil, Faisal, and Eatzaz Ahmad. 2011. Income and price elasticities of electricity demand: aggregate and sector-wise analyses. *Energy Policy* 39 (9): 5519–27.

Jenkins, Jesse, Ted Nordhaus, and Michael Shellenberger. 2011. *Energy emergence: rebound and backfire as emergent phenomena.* Oakland, CA: Breakthorugh Institute.

Jessoe, Katrina, and David Rapson. 2014. Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review* 104 (4): 1417–38.

Jevons, William Stanley. 1865. *The coal question; an enquiry concerning the progress of the Nation, and the probable exhaustion of our coal-mines.* 2nd ed. London,: Macmillan.

Li, Shanjun, Joshua Linn, and Erich Muehlegger. 2014. Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy* 6 (4): 302–42.

Lin, C.-Y. Cynthia, and Jieyin Zeng. 2013. The elasticity of demand for gasoline in China. *Energy Policy* 59 (C): 189–97.

Linn, Joshua. 2013. The rebound effect for passenger vehicles. Working Paper, RFF.

Nahata, Babu, Alexei Izyumov, Vladimir Busygin, and Anna Mishura. 2007. Application of the Ramsey model in transition economy: A Russian case study. *Energy Economics* 29 (1): 105–25.

Ramanathan, Raj. 1999. Short- and long-run elasticities of gasoline demand in India: An empirical analysis using co-integration techniques. *Energy Economics* 21 (4): 321–30.

Ramey, Valerie. 2011. Can government purchases stimulate the economy. *Journal of Economic Literature* 49 (3): 673–85.

Saunders, Harry. 1992. The Khazzoom-Brookes postulate and neoclassical growth. *Energy Journal* 13 (4): 131–48.

———. 2013. Historical Evidence for energy consumption rebound in 30 US sectors and a toolkit for rebound analysts. *Technological Forecasting and Social Change* 80 (7): 1317–30.

Sene, Seydina Ousmane. 2012. Estimating the demand for gasoline in developing countries: Senegal. *Energy Economics* 34 (1): 189–94.

Small, Kenneth A., and Kurt van Dender. 2007. Fuel efficiency and motor vehicle travel: The declining rebound effect. *Energy Journal* 28 (1): 25–51.

Sorrell, Steve. 2007. *The rebound effect: An assessment of the evidence for economy-wide energy savings from improved energy efficiency.* UK Energy Research Center Report.

Sorrell, Steve, and John Dimitropoulos. 2008. The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics* 65 (3): 636–49.

Thiesen, J., T. Christensen, T. Kristensen, R. Andersen, B. Brunoe, T. Gregersen, M. Thrane, and B. Weidema. 2008. Rebound effects of price differences. *International Journal of Life Cycle Assessment* 13 (2): 104–14.

Thomas, Brinda A., and Ines L. Azevedo. 2013. Estimating direct and indirect rebound effects for U.S. households with input-output analysis. Part 2: Simulation. *Ecological Economics* 86 (C): 188–98.

Tsao, Jeffrey, Harry Saunders, J. Randall Creighton, Michael Coltrin, and Jerry Simmons. 2010. Solid-state lighting: An Energy-economics perspective. *Journal of Physics D: Applied Physics* 43: 1–17.

Turner, Karen. 2009. Negative rebound and disinvestment effects in response to an improvement in the UK economy. *Energy Economics* 31 (5): 648–66.

van den Bergh, Jeroen. 2011. Energy conservation more effective with rebound policy. *Environmental and Resource Economics* 48 (1): 43–58.

West, Sarah. 2004. Distributional effects of alternative vehicle pollution control technologies. *Journal of Public Economics* 88 (3–4): 735–57.

Wolfram, Catherine, Orie Shelef, and Paul Gertler. 2012. How will energy demand develop in the developing world? *Journal of Economic Perspectives* 26 (1): 119–38.

Zein-Elabdin, Eiman. 1997. Improved stoves in Sub-Saharan Africa: The case of the Sudan. *Energy Economics* 19 (4): 465–75.

Ziramba, Emmanuel. 2008. The demand for residential electricity in South Africa. *Energy Policy* 36 (9): 3460–66.

Environmental Science & Technology

Policy Analysis

pubs.acs.org/est

Compliance by Design: Influence of Acceleration Trade-offs on CO_2 Emissions and Costs of Fuel Economy and Greenhouse Gas Regulations

Kate S. Whitefoot,[*,†] Meredith L. Fowlie,[‡] and Steven J. Skerlos[§]

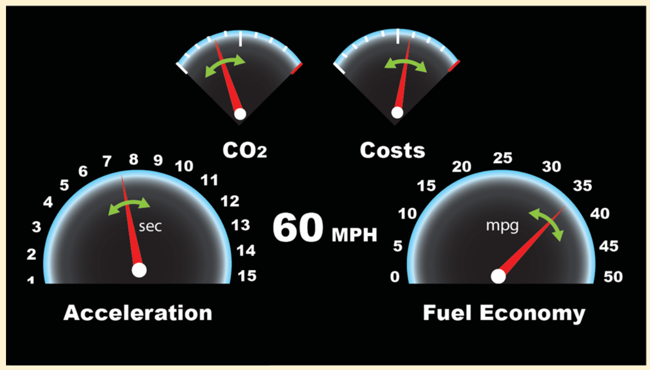
[†]Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, United States

[‡]University of California 207 Giannini Hall #3310, Berkeley, California 94720, United States

[§]University of Michigan, 500 S. State Street, Ann Arbor, Michigan 48109, United States

Supporting Information

ABSTRACT: The ability of automakers to improve the fuel economy of vehicles using engineering design modifications that compromise other performance attributes, such as acceleration, is not currently considered when setting fuel economy and greenhouse-gas emission standards for passenger cars and light trucks. We examine the role of these design trade-offs by simulating automaker responses to recently reformed vehicle standards with and without the ability to adjust acceleration performance. Results indicate that acceleration trade-offs can be important in two respects: (1) they can reduce the compliance costs of the standards, and (2) they can significantly reduce emissions associated with a particular level of the standards by mitigating incentives to shift sales toward larger vehicles and light trucks relative to passenger cars. We contrast simulation-based results with observed changes in vehicle attributes under the reformed standards. We find evidence that is consistent with firms using acceleration trade-offs to achieve compliance. Taken together, our analysis suggests that acceleration trade-offs play a role in automaker compliance strategies with potentially large implications for both compliance costs and emissions.



INTRODUCTION

The U.S. Corporate Average Fuel Economy (CAFE) and Greenhouse Gas (GHG) standards, issued by the National Highway and Traffic Safety Administration (NHTSA) and Environmental Protection Agency (EPA) are the principal means of reducing GHG emissions of light-duty vehicles in the United States. A significant reform of these standards occurred after the passage of the Energy Independence and Security Act (EISA) in 2007. The reformed standards do not set a fixed level of fuel economy or GHG emissions that must be met. Instead, the standards for each automaker are based on the sizes of the vehicles they produce (specifically, the vehicle's *footprint*, defined as the wheelbase multiplied by the track width) and various credits they can receive (e.g., alternative-fuel vehicle credits). The first phase of these reformed standards were enforced between 2011 and 2016. The agencies have since issued standards for 2017−2021 and are evaluating the costs and benefits of the policy to inform the final standards through 2025.

NHTSA is required to set the standards at the "maximum feasible" level, considering "technological feasibility, economic practicability, the effect of other motor vehicle standards of the government on fuel economy, and the need of the United States to conserve energy."[1] The agencies have met this requirement by determining the costs and benefits of adopting various technologies that reduce fuel consumption and GHG emissions while maintaining or improving the performance of other vehicle attributes, most notably acceleration time.[2] This cost-benefit analysis informs the standard-setting along with other considerations, such as harmonization with state GHG regulations.[3]

One advantage of the agencies' analytical approach is that it guarantees the standards can be met using available technologies, assuming vehicle demand does not change.[2] Still, nothing restricts automakers to respond to the standards the way the agencies' model predicts. Automakers have multiple compliance options available to them, and presumably choose the combination of strategies that minimize their compliance costs. Policy analyses that do not account for the full suite of compliance options may significantly overestimate compliance costs and produce misleading estimates of emission reductions.

Received: July 21, 2017
Accepted: August 21, 2017
Published: August 21, 2017

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

In addition to implementing various technology features, other possible responses to the policy include (1) trading off vehicle performance attributes (such as acceleration performance) to improve fuel economy,[4−6] (2) taking advantage of various credit provisions,[7,8] (3) adjusting prices to shift sales to vehicles that exceed their fuel economy target,[9−11] (4) increasing vehicle footprint (thereby decreasing the stringency of their fuel economy and GHG targets),[12] and (5) violating the standards and paying fines to NHTSA and civil penalties to EPA.[13−15] Previous studies have examined the influence of the latter four of these alternative strategies on fuel consumption and/or costs.[7,9−12] Whether firms have incentives to trade off acceleration performance and fuel economy in response to the reformed policy, however, has not been examined in depth.

In this paper, we investigate the role that engineering design trade-offs between acceleration performance and fuel economy can play in automakers' response to the reformed standards. To do this, we nest a flexible approximation (also called a surrogate model) of engineering design trade-offs generated from physics-based vehicle performance simulations within an economic equilibrium model of the automotive market. We then simulate the engineering design and pricing decisions of profit-maximizing firms responding to the 2014 standards with and without the ability to trade off acceleration performance.

Our analysis focuses on the compliance options that automakers can use over the "medium run", namely fuel-efficiency technologies and design trade-offs that can be implemented in the first few (i.e., 1−6) years after the regulations are announced. In order to be consistent with the agencies' approach, we do not account for design changes to vehicle footprint and compliance options that take longer production planning lead times, such as converting a significant percentage of their fleet to electric vehicles.[13] However, we also find that our conclusions are robust to relaxing the technology assumptions.

Unlike the agencies' analysis, which assumes vehicle-specific demand is fixed, our model allows demand to respond to policy-induced changes in vehicle prices and attributes. This demand response is important to consider when assessing the significance of acceleration trade-offs. In contrast with fuel-efficiency technologies that increase vehicle production costs, the primary costs to automakers of compromising acceleration performance are lost profits due to reduced demand and/or lower markups necessary to achieve a particular level of demand.

Acceleration trade-offs can lower the compliance costs associated with the regulation in three related ways. First, an automaker may find it relatively more profitable to compromise the acceleration performance of its vehicles (to improve fuel economy) rather than incorporating additional costly fuel-saving technologies or changing prices to shift demand to more fuel-efficient vehicles. Second, automakers may prefer to use acceleration trade-offs in combination with technology features in some or all of their vehicles so that fuel economy improves as well as acceleration performance. Third, if the regulation induces worse acceleration performance in some vehicles, competition for consumers who value acceleration will be reduced. This may cause some automakers to *improve* the acceleration performance of certain vehicles (in order to attract these consumers) at the expense of fuel economy, while simultaneously improving the fuel economy of other vehicles enough to comply with the standards.

This paper contributes to a growing body of literature that examines the economic and environmental impacts of fuel economy and GHG standards. Recent research finds that manufacturers can use a variety of loopholes and other compliance mechanisms that relax the stringency of the standards, leading to higher emissions.[7,8,12,15,16] If acceleration trade-offs offer a relatively cost-effective means of complying with the standards, automakers' incentives to exploit these mechanisms that relax the stringency of the standards will be reduced.

Our work also begins to bridge a gap between the engineering design and economics literatures examining firms' optimal product design and pricing decisions. The approach we take is designed to leverage the relative strengths of methods in each field. Recent work in the economics literature uses bundles of attributes observed in the marketplace to econometrically estimate engineering trade-offs between energy efficiency and other product attributes.[5,11,17,18] The most closely related example is Klier and Linn (2012), who examine the influence of trade-offs between fuel economy and engine power in the context of the prereform CAFE standards. One limitation of this approach is that many combinations of product attributes are not observed in the marketplace, but are technologically feasible and potentially optimal under future policy scenarios. A second concern is that correlations between attributes of interest (e.g., energy efficiency) and attributes that are difficult to quantify or otherwise unobservable in historical data (e.g., vehicle shape) can make it difficult to identify attribute trade-offs econometrically. The physics-based engineering simulations we use to characterize design trade-offs can identify technologically possible combinations of attributes that have yet to manifest in existing product designs. This approach also allows us to identify trade-offs independently of unobserved product attributes.

The engineering design literature, on the other hand, develops detailed models of the trade-offs among product attributes based on physics.[19−21] In this literature, it is common to determine a particular firm's choices of engineering design variables and prices that maximize the firm's profits.[22−24] With a few notable exceptions,[25−29] however, this body of research generally ignores the strategic nature of competing firms' price and design decisions. The studies that do account for competitor design and pricing decisions are focused on relatively simple examples with ten or fewer products in the market and identical design trade-offs and costs for all firms. We extend this literature by nesting an engineering-design model of heterogeneous firms producing many product variants (a total of 471 distinct vehicle models and engine options) in an economic equilibrium model that captures the strategic competition between automakers. This extension is significant because the strategic interactions between competing firms and the industry structure affects firms' profit-optimal designs and prices,[25,29] and therefore resulting emissions and costs.

■ MATERIALS AND METHODS

To capture the trade-offs between acceleration performance and fuel economy, we implement thousands of vehicle performance simulations over a range of feasible vehicle design configurations using an engineering simulation software package (AVL Cruise) that is used by the automotive industry to support the powertrain development process. To incorporate these simulated data in our model in a tractable way, we estimate a flexible approximation of the relationships among

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

vehicle performance attributes and production costs. These estimated relationships are then nested within an oligopolistic equilibrium model of the automotive market.

On the supply side, we include the 18 automakers that comprise 97% of the U.S. market. We assume each firm chooses prices and design variables for each of their vehicle models and engine options (e.g., the Toyota Camry with a 2.5 L engine and with a 3.5 L engine) to maximize profits. More specifically, we allow automakers to adjust fuel consumption (measured as gallons of fuel consumed per 100 miles) and acceleration (measured as the time in seconds to accelerate from 0 to 60 mph) by modifying powertrain tuning variables and technology features that can be changed in the medium run during vehicle redesign. We hold fixed the vehicle design parameters that are determined in earlier stages of the vehicle development process (see Supporting Information (SI) S1.1 for details). Longer-run design parameters include vehicle segment (e.g., midsize sedan), the powertrain architecture (e.g., conventional gasoline, hybrid, or diesel), and key internal and external dimensions.[30−32]

On the demand side, a random-coefficient logit discrete choice model is estimated using household-level data on vehicle purchase decisions. Taken together, the supply and demand-side models can be used to simulate how automakers' profit-maximizing choices of vehicle designs and prices change in response to the 2014 standards, and the resulting impact on emissions and costs in equilibrium. To evaluate how acceleration trade-offs affect these outcomes, we generate two sets of simulations: (1) a model where automakers can adjust acceleration performance and fuel consumption, and (2) a more restricted model where acceleration performance is held fixed for all vehicles.

We choose 2006 as the reference year for consumer preferences and "baseline" vehicle designs to which automakers can add technology options and adjust powertrain tuning variables. This was the year immediately preceding the passage of EISA. After this year, automakers presumably began to plan their compliance strategies, and in some cases, implement design changes to earn early compliance credits.

Engineering Design Trade-offs. We make a conceptual distinction in our modeling framework between two types of engineering design modifications that automakers can use to change the fuel economy of their vehicles in the medium-run. *Powertrain tuning variables* (e.g., the final drive ratio) can be adjusted to favor fuel economy over acceleration performance or vice versa and have negligible influence on production costs or lower these costs. *Technology features* can be incorporated into a vehicle at an extra cost to improve fuel economy. Examples of technology features include high-efficiency alternators, low resistance tires, and low-friction materials in the engine. Many (although not all) of these technology features improve acceleration performance in addition to fuel economy.

As we discuss below, our model of the vehicle development process is not comprehensive. Because of simulation and data constraints, we do not account for all powertrain tuning variables and technology features automakers have at their disposal in the medium-run to increase fuel economy (e.g., changing the number of transmission gear ratios). If excluded powertrain tuning variables or technology features are less cost-effective to change than those explicitly accounted for, omitting them will be inconsequential. If any of the omitted powertrain tuning variables are more cost-effective, our results represent lower bounds of the impact that design trade-offs can have on emissions and costs. However, if omitted technology features are more cost-effective than those we include, the influence of acceleration trade-offs would be overestimated. To assess the robustness of our findings to the set of technology features considered, we conduct sensitivity tests of our results to extending the technology improvements possible and lowering technology costs.

Our modeling of vehicle design trade-offs begins with the construction of "bundles" of design variables specific to each vehicle segment, s, indexed $b = 1 \cdots B$. Each bundle is comprised of a set of powertrain tuning variables, $\boldsymbol{x}_s$, and technology features, t_s, that firms are able to adjust in our equilibrium model, as well as fixed design parameters, $\tilde{x}_s$, which firms cannot change. In the model, there are two powertrain tuning variables that can be manipulated to trade off acceleration performance for improved fuel economy: engine displacement size and the final drive gear ratio in the transmission. Fixed design parameters consist of vehicle segment, baseline curbweight (i.e., the weight of the vehicle without any passengers or cargo and without substituting existing materials for lightweight materials), gradeability (i.e., the steepest hill a vehicle can climb maintaining a particular speed), and towing requirements. Our classification of vehicle parameters as adjustable or fixed is based on the structure of the vehicle development process and manipulability of these parameters over the medium run as described in detail in SI S1.1. Technology features are taken from NHTSA's analysis of available fuel-saving technologies based on independent studies and information from automotive manufacturers, researchers, and consultants (SI Table S2).[33]

We use the vehicle performance simulation package AVL Cruise to calculate the fuel consumption per 100 miles (fuelcons) and 0−60 mph acceleration time (acc) of a particular vehicle design conditional on a specified bundle of design parameters, b. We generate almost 30 000 sets of simulation results, each representing the fuel consumption and acceleration performance corresponding to the bundle of design parameter inputs, which are varied at small increments. Additional details of the vehicle simulations are discussed in SI S1.3.

The relationship between adjustable powertrain tuning variables and production costs is taken from Michalek et al., who estimate the relationship using data from automotive manufacturers and wholesale and rebuilt engine suppliers.[27] Production costs associated with the addition of specific technology features are taken from NHTSA's analysis (SI Table S2), which were used in cost-benefit analyses of the regulations.[33] NHTSA collected these cost data from vehicle tear-down studies, confidential manufacturer information, and independent studies. Cost reductions due to learning in the time between the announcement of the reformed regulations and their implementation are incorporated into the agencies' estimates (see SI S1.6−1.8 for a detailed description of the data). Similar to the agencies' approach, we assume that all changes to vehicle designs occur during regularly scheduled product redesign cycles and so do not incur additional costs that would be associated with modifying the medium-run vehicle design variables in later stages of the development process.[34]

Because changes to the final drive ratio negligibly influence production costs, for any chosen values of acceleration performance and technology features, there is only one choice of $\boldsymbol{x}_s$ that minimizes the production costs, c, associated with a

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

given level of fuel consumption (see SI S1.4 for a detailed explanation). The engineering design trade-offs we model can thus be summarized by a system of two equations representing the efficiency frontiers (called Pareto frontiers in the engineering design literature) of fuel consumption and production costs for a particular vehicle design as a function of its acceleration performance and technology features, conditional on fixed design parameters (derivations are provided in SI S1.4 and S1.6): $\text{fuelcons}_{\text{sb}} = h_{1s}(\text{acc}_{\text{sb}}, t_s; \tilde{x}_s)$, $c_{\text{sb}} = h_{2s}(\text{acc}_{\text{sb}}, t_s; \tilde{x}_s)$.

While we could in principle specify the structure of these two functions and estimate the parameters separately for all possible combinations of technology features, in practice it is computationally infeasible to explicitly incorporate this large number of discrete technology combinations in our equilibrium simulations. For the purpose of tractability, we approximate the set of cost-effective technology feature combinations with a single continuous variable, tech. The tech variable takes on a value between zero (the baseline case) and the maximum number of cost-effective combinations of technology features for each vehicle segment, with each value mapping to a specific combination of technology features. These technology combinations are ordered by decreasing fuel consumption for the same acceleration time, which is also increasing in cost. Therefore, a higher value of tech corresponds to a lower fuel-consumption and higher cost vehicle conditional on 0−60 mph acceleration time.

Several parametric specifications of the fuel consumption and cost functions were estimated using the vehicle simulation and production cost data. The following specifications performed the best under the Akaike Information Criterion:

$$\begin{aligned}\text{fuelcons}_{\text{sb}} &= \kappa_{1s} + \kappa_{2s}\text{e}^{-\text{acc}_{\text{sb}}} + \kappa_{3s}wt_{\text{sb}} + \kappa_{4s}wt_{\text{sb}}\cdot\text{acc}_{\text{sb}} \\ &\quad + \kappa_{5s}\text{tech}_{\text{sb}} + \kappa_{6s}\text{tech}_{\text{sb}}\cdot\text{acc}_{\text{sb}} + \epsilon_{\text{sb}} \qquad (1)\end{aligned}$$

$$\begin{aligned}c_{\text{sb}} &= \sigma_{1s} + \sigma_{2s}\text{e}^{-\text{acc}_{\text{sb}}} + \sigma_{3s}wt_{\text{b}} + \sigma_{4s}wt_{\text{b}}\cdot\text{acc}_{\text{sb}} + \sigma_{5s}\text{tech}_{\text{s}} \\ &\quad + \epsilon_{\text{sb}} \qquad (2)\end{aligned}$$

where fuelcons, c, acc, and wt, are the fuel consumption, marginal production costs, 0−60 mph acceleration time, and the curbweight of a vehicle in segment s with bundle of design variables b. We show in SI S1.9 that these particular specifications preserve important relationships between fuel consumption, acceleration performance, technology features, and costs from the underlying vehicle performance simulations and cost data.

Estimated values of the parameters in eqs 1 and 2 are reported in SI Tables S3 and S4. The models fit the data in each segment reasonably well (R^2 = 0.81−91) with the exception of the two-seater segment (R^2 = 0.67 for fuel consumption and 0.75 for costs). However, this segment comprises less than 1% of vehicle sales so the poorer fit should not significantly affect the policy simulation results.

Demand Model. Following Train and Winston (2007), we model consumer vehicle choices using a random-coefficient logit model estimated using data on consumer-level choices and vehicle attributes. The utility consumer n derives from vehicle model and engine option i can be decomposed into four components:

$$u_{ni} = \delta_i + \sum_{\mathbf{kr}} a_{ik}z_{nr}\beta_{kr} + \sum_{\mathbf{k}} a_{ik}\nu_{nk}\mu_k + \epsilon_{ni} \qquad (3)$$

The first component, δ_i, captures the average utility across consumers for a specific vehicle model and engine option. $\delta_i = \sum_k a_{ik}\overline{\beta_k} + \xi_i$ where each a_{ik} is an observable vehicle attribute, such as price and fuel economy, $\overline{\beta}_k$ is the coefficient for the attribute, and ξ_i captures the utility of attributes valued by the consumer but not observed in the data (e.g., interior materials). The second component represents the portion of utility for vehicle attributes that varies systematically with observed consumer characteristics, $\boldsymbol{z}_n$. The third component captures the effects of interactions between vehicle attributes and consumer characteristics we cannot observe. This allows for random variation in consumer preferences for specific vehicle attributes, $\boldsymbol{\mu}$, which are assumed to be normally distributed.

The fourth term, ϵ_{nj}, in eq 3 captures idiosyncratic individual preferences. We invoke the standard assumption that these errors have an i.i.d. Type I extreme value distribution. This assumption yields the following functional form for the vehicle-choice-share probabilities, P_{ni}, conditional on $\boldsymbol{z}_n$, $\boldsymbol{\nu}_n$, and the parameters to be estimated, $\boldsymbol{\theta}$.

$$\begin{aligned}P_{ni} = \Pr(y_n = i|z_n, \nu_n, \boldsymbol{\theta}) &= \frac{\exp(\delta_i + \sum_{\mathbf{kr}} a_{ik}z_{nr}\beta_{kr} + \sum_k a_{ik}\nu_{nk}\mu_k)}{1 + \sum_j \exp(\delta_j + \sum_{kr} a_{jk}z_{nr}\beta_{kr} + \sum_k a_{jk}\nu_{nk}\mu_k)} \\ &\equiv \frac{\exp(u_{ni})}{1 + \sum_j \exp(u_{nj})} \qquad (4)\end{aligned}$$

The predicted sales of vehicle i is $M\sum_n P_{\text{in}} \equiv q_i$ where M is the market size. This utility formulation is extended to include consumers' ranked choices when available (see SI S3.2).

Because unobserved vehicle attributes that consumers value, such as interior materials, acoustic performance, and electronic accessories, are likely to be correlated with the vehicle attributes of primary interest (namely, price, fuel economy, and acceleration performance), estimating eq 4 for $\overline{\beta}_k$ directly will likely yield biased estimates. This well-documented endogeneity problem is typically addressed using an instrumental variables (IVs) strategy.[35−38] It has become standard to use functions of nonprice attributes, $\boldsymbol{w}$, including horsepower and fuel economy, as IVs for endogenous attributes.[35−39] This strategy is predicated on an exclusion restriction that requires the IVs to be exogenous such that $\mathbb{E}[\xi_i|\mathbf{w}] = 0$. Our study is motivated by the observation that automakers can modify vehicle attributes such as fuel economy and horsepower in the medium-run. Thus, in contrast to earlier studies, we use only those vehicle attributes that are determined by longer run product-planning schedules as IVs for price, fuel economy, and acceleration performance. Specifically, we use the moments of vehicle dimensions of same-manufacturer vehicles and different-manufacturer vehicles, powertrain architecture (e.g., hybrid, diesel, conventional gasoline), and drive type (e.g., all wheel drive). This identification strategy is discussed in more detail in SI S3.3.

Two sources of data are used to estimate the demand model: a detailed household-level survey conducted by Maritz Research in 2006, and vehicle characteristic data available from Chrome Systems Inc. SI S3.1 describes these data and reports the estimated parameters in SI Tables S8 and S11. We perform random initial value tests and verify that the algorithm converges to the same solution.

Automotive Oligopoly Model. To model firms' product pricing and design decisions, we nest the engineering design and demand models summarized by eqs 1−4 within a differentiated product oligopoly model. We assume that firms choose the prices, acceleration performance, and levels of

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

technology features of all the vehicle models and engine options they produce to maximize profits, π, according to the following formulation.

$$\max_{p_j, \mathrm{acc}_j, \mathrm{tech}_j \forall j} \pi = \sum_j q_j(p_j - c_j)$$

$$\text{subject to } \mathrm{CAFE}_l^{\mathrm{TARGET}} - \mathrm{CAFE}_l - \mathrm{credit}_l \le 0 \ \forall \ l$$

$$\text{where } q_j = g(p_j, \mathrm{fuelcons}_j, \mathrm{acc}_j; \tilde{\mathrm{x}}_j)$$

$$\mathrm{fuelcons}_j = \mathrm{h}_1(\mathrm{acc}_j, \mathrm{tech}_j; \tilde{\mathrm{x}}_j)$$

$$c_j = \mathrm{h}_2(\mathrm{acc}_j, \mathrm{tech}_j; \tilde{x}_j) \tag{5}$$

The variables q_j, p_j, and c_j are, respectively, the quantity demanded, price, and marginal cost associated with vehicle model and engine option j. The standards are represented as a constraint for each vehicle class l (i.e., passenger cars and light trucks). We define CAFE_l to be the harmonic sales-weighted average fuel economy of all vehicles in class l that the firm produces, which must equal or exceed the firm's CAFE target for that vehicle class, $\mathrm{CAFE}_l^{\mathrm{TARGET}}$, within allowable fuel-economy credit provisions defined by the regulations, credit_l. Excluding differences between noncompliance penalties and the credits automakers can earn under the CAFE and GHG standards (i.e., AFV and off-cycle credits), the standards are equivalent. Therefore, in the case where firms meet the standards without the use of these credits, both standards can be represented by the single constraint for each vehicle class in eq 5. We repeat the simulations under alternative assumptions to explore scenarios under which automakers can earn additional credits under the CAFE and GHG standards (see SI S4.3).

For the main specification presented in this paper, we allow all firms to trade credits between their passenger car and light truck fleets but we constrain firms to comply with the 2014 standards without further credit provisions. We use our oligopoly model to simulate the effects of replacing the unreformed 2006 standards with the 2014 reformed standards with and without the consideration of acceleration trade-offs. We use 2006, which just predates the policy reform, as a baseline against which we determine emission reductions, changes in vehicle attributes, and producer and consumer costs. To avoid confounding the effects of the policy reform with our modeling assumptions (including any model misspecification and the omission of some credit provisions) as well as the exogenous reduction of technology costs over time, we use *simulated* partial equilibrium outcomes under the 2006 standards as our baseline rather than observed data. A comparison of the simulated baseline outcomes with observed attributes is provided in SI S4.1. In order to build confidence in our simulations, we perform out-of-sample predictions of sales-weighted average fuel economy and acceleration performance in the years between 2006 and 2014 and compare them to observed values in these years. We find that the simulations predict observe values within 3% for each year (see SI S4.5).

SIMULATION RESULTS

The model is used to simulate a series of vehicle-specific equilibrium outcomes: fuel economy, acceleration performance, technology features, prices, production costs, and vehicle sales. These simulated outcomes are used to calculate total use-phase GHG emission reductions over the lifetime of the vehicles and producer and consumer costs resulting from replacing the 2006 standards with the 2014 standards. GHG emissions are calculated assuming passenger cars and light trucks are respectively driven 195 000 and 226 000 miles over their lifetime in the baseline with a rebound effect of 10.3%.[40,41] Producer costs are measured in terms of profit losses relative to the baseline. Consumer costs are measured in terms of consumer surplus losses calculated by equivalent variation, or the amount that a consumer would need to be paid to realize the same amount of utility. We determine the compliance costs of the policy in terms of the sum of profit losses and consumer surplus losses (hereafter, social surplus losses) per ton of emissions reduced. We stop short of a comprehensive measure of the societal benefits (e.g., improved air quality) associated with reduced fuel consumption and GHG emissions in these calculations.

In addition to assessing the extent that acceleration trade-offs influence GHG emissions and social surplus, we investigate two "offsetting" effects that play a role in determining the net effect of the reformed standards on aggregate emissions. The first relates to the differences in stringency between the passenger-car and light-truck standards. If the market share of light trucks rises relative to that of passenger cars, GHG emissions will be higher. The second relates to the fact that the standards are size-based. Firms can reduce the stringency of the standards by shifting sales toward larger passenger cars and light trucks.

Table 1 summarizes simulation results for two scenarios: (1) modeling trade-offs between fuel economy and acceleration performance, and (2) excluding these trade-offs. In the simulation that include design trade-offs, we see significant compromises in acceleration performance. Average 0–60 mph acceleration time increases 0.7 s or approximately 8% (an increase in acceleration time means acceleration performance is worse). Notably, the large majority of this change comes from the design response versus changes in sales composition. In the simulations that shut off the design trade-offs, we see a relatively small increase in acceleration time, which is driven entirely by changes in sales composition.

Results indicate that, when acceleration trade-offs are considered, GHG emission reductions increase from 19 to 77

Table 1. Simulation Results of the Impact of Replacing the 2006 Standards with the 2014 Standards in Simulations Excluding and Including Acceleration (acc) Trade-Offs; Both Scenarios Account for Price Changes and Adoption of Fuel-Saving Technology Features

	no acc trade-offs	with acc trade-offs
Emissions and Social Surplus Results		
change in CO_2 emissions (million metric tons)	−18.8	−76.5
change in consumer surplus (billions 2014 USD)	−\$12.8	−\$7.4
change in producer profits (billions 2014 USD)	−\$0.2	−\$0.1
Aggregate Vehicle Attribute Results		
change in sales-weighted average fuel economy (mpg)	+0.7	+2.6
change in sales-weighted average acceleration (s)	+0.1	+0.7
change in sales-weighted average footprint (sq ft)	+1.2	+0.1
change in share of light trucks	+3.5%	+2.7%

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

million metric tons. There are two key reasons for this that are related to changes in the composition of new cars sold. First, the market share of light trucks increases more when acceleration trade-offs are shut off. Second, the sales-weighted average vehicle footprint increases by 1.0 sq. ft (0.09 m^2) when acceleration trade-offs are excluded, whereas it remains approximately the same when they are included. Recall that we do not allow firms to change the footprint of their vehicles in our simulations, so this increase in size is due to price changes that shift demand to larger passenger cars and light trucks (see Whitefoot and Skerlos[12] for an analysis of size increases when footprint-design changes are possible).

Social surplus is also significantly impacted by acceleration trade-offs. Policy-induced consumer surplus losses decrease from \$12.8 billion or approximately \$790 per consumer to \$7.4 billion or approximately \$460 per consumer when the trade-offs are included. Total profit losses are reduced from \$200 million to \$100 million, which should be considered upper bounds because we do not account for all compliance flexibilities in the regulations (e.g., banking and borrowing of credits). Total social surplus losses when attribute trade-offs are excluded are comparable to results in Klier and Linn's (2012) study of the prereform regulations after adjusting for the stringency of the reformed standards (see S4.4 for details). The change in social surplus when attribute trade-offs are included, however, is smaller than that reported in Klier and Linn. This is most likely due to a combination of two factors. First, our estimates of attribute trade-offs using physics-based vehicle simulations imply that fuel consumption can be reduced with smaller adjustments in acceleration performance than econometric estimates that may conflate trade-offs with unobserved vehicle attributes correlated with fuel economy and acceleration (see S2). Second, the reformed policy differs from past regulations in several important ways that reduce costs for compliant firms (e.g., reducing leakage by enforcing tough penalties for firms that violate the GHG standards). Similar to most prior studies, we find that the vast majority of the costs of the policy are passed on to consumers.

The simulated average impacts on vehicle attributes mask significant heterogeneity across vehicles. Figure 1 shows the policy-induced changes in sales-weighted average fuel economy and acceleration times and the spread between the 10th and 90th quantiles. As the figure illustrates, there are mostly increases, but also some notable decreases, in these attributes

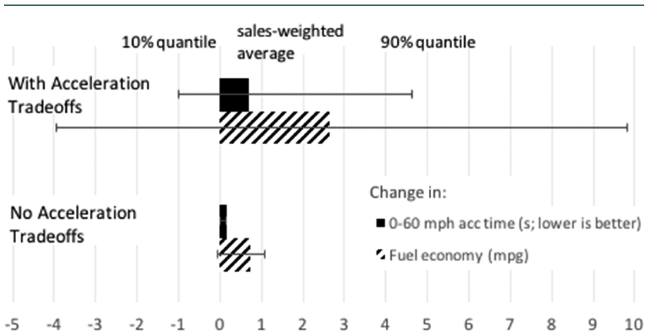


Figure 1. Simulation results of changes in vehicle attributes in response to the reformed standards when trade-offs between fuel economy and acceleration (acc) are considered and when they are excluded from the analysis. The sales-weighted averages for fuel economy are harmonic averages following the policy, whereas arithmetic averages are used for acceleration.

when acceleration trade-offs are included. Recall that automakers may reduce the fuel economy of some vehicles in favor of acceleration performance to attract consumers willing to pay for superior acceleration performance. Our simulation results show that for 25% of vehicles, firms choose to reduce fuel economy in order to improve acceleration performance. For 57% of vehicles, automakers rely on compromising acceleration performance rather than relying on fuel-saving technology features to improve fuel economy, and for 17% they use a combination of acceleration trade-offs and technology features. Less than 1% have no change in acceleration performance. We also find heterogeneity across automakers. Some firms rely on acceleration trade-offs to comply with the standards to a much greater extent than others.

We also conduct sensitivity tests that examine the effect of varying the estimates of technology feature costs and consumer willingness-to-pay for fuel economy. Results are summarized in Table 2. When consumers are willing to pay more for

Table 2. Influence of Acceleration Trade-Offs on Simulation Results of the 2014 Policy Outcomes under Alternate Specifications[a]

	change in CO_2 emissions (million metric tons)	change in compliance costs (billion 2014 USD)
main specification	−58	−5.5
Sensitivity Tests on Main Specification		
willingness to pay for fuel economy 35% higher	−32	−2.2
cost of tech features 25% lower	−48	−4.7
maximum tech 10% higher	−54	−3.7

[a]The table presents the *difference* between the estimates produced from the simulations where acceleration trade-offs are included with the estimates produced by the simulations that exclude these trade-offs.

improvements in fuel economy, the influence of acceleration trade-offs on emissions and compliance costs is lower than in the main simulation specification. The change in GHG emission reductions due to acceleration trade-offs drops from 58 million metric tons to 32 and reductions in compliance costs drop from 5.5 billion to 2.2. This occurs because the standards are effectively less stringent so that the benefits of using acceleration trade-offs as an additional compliance strategy is smaller (although still substantial). Intuitively, acceleration trade-offs also have a somewhat smaller impact on emissions and compliance costs when fuel-saving technology costs are lower and when the upper bound of the tech variable is relaxed.

■ LONGITUDINAL REGRESSION OF OBSERVED ACCELERATION

The simulation results summarized above predict how the reformed policy affects acceleration trade-offs, emissions, and compliance costs conditional on modeling assumptions and holding other confounding factors (such as fuel prices) constant. These simulations are based entirely on data that was available before the policy change in order to be consistent with the type of analysis agencies could perform when assessing the impact of future policy options.

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307−10315

As a check on the simulation results, we also examine acceleration trade-offs using a completely different method: a longitudinal regression analysis of the acceleration performance we observe in the new vehicle market before and after the reformed standards took effect. We use data on sales-weighted attributes collected by EPA for 1976–2014 vehicles.[42] These data are recorded at the level of firm-year for each vehicle class. Although the reformed standards did not apply until 2011, automakers could earn credits for earlier action, which they could use to comply with the standards once they took effect.[43] For this reason, we look for evidence of policy-induced design changes as early as 2007 (after EISA was passed).

Empirically estimating the causal effect of the policy reform on vehicle attributes is difficult because there are many time-varying factors that could influence vehicle design choices. Potentially confounding factors include exogenous technological change, rising gasoline prices, and evolving consumer preferences. In order to isolate the effect of the reformed standards on vehicle design choices as best we can with the available data, we include several controls for these time-varying factors in our analysis.

We use 30 years of data prior to the announcement of the reformed policy to analyze trends in acceleration performance over time. The following equation serves as the foundation for our empirical analysis:

$$\mathrm{acc}_{it} = \alpha + \delta(t) + \beta' X_{it} + \gamma_1 \mathrm{D1}_t + \gamma_2 \mathrm{D2}_t + \varepsilon_{it} \quad (6)$$

where i indexes manufacturing firms and t indexes time (measured in years). The $\delta(t)$ function models acceleration performance as a function of time. X_{it} captures time varying determinants of acceleration performance such as gasoline prices. D1 and D2 are policy indicators that equal one one after MY2006 and MY2010, respectively, and zero before. Including these binary policy indicators allows a level shift in acceleration performance trends after the policy takes effect. We also estimate a linear spline function which allows the slope of the acceleration performance trajectory to change as firms begin to comply with the policy. In the spline specifications, binary indicators in eq 25 are replaced with $\mathrm{B1}_t = t\text{-}2006$ and $\mathrm{B2}_t = t\text{-}2010$.

Results are summarized in Table 3 (additional specifications are described in the SI). Relative to the trends and relationships observed prior to the reformed standards, we find that the rate of improvement in acceleration performance slowed after the policy reform was announced and slowed further once the policy took effect. These policy variables are jointly significant. The preferred specifications are (2) and (4), which condition on real gasoline prices. These estimated coefficients can be used to impute an effect of the policy on sales-weighted average 0–60 mph acceleration time. The table reports these imputed effects which range from 0.63–1.10 s slower. For the preferred specifications, the estimated effects of the policy on average acceleration are remarkably similar to our simulation-based estimate of 0.7 s.

Analyzing these same data at the firm-level reveals substantial heterogeneity in patterns of acceleration performance across manufacturers. For each firm and vehicle type (i.e., passenger car or light truck), we construct the counterfactual trajectory of acceleration performance by extrapolating prepolicy acceleration trends controlling for time-varying factors. Observed acceleration following the introduction of the reformed standards underperforms relative to this counterfactual for most firms. For some firms, however, we estimate improvements in acceleration performance among passenger cars (Kia) and trucks (Chrysler, Ford, and Mercedes-Benz). These firm-level estimates are summarized in SI Table S14. While the firm-level heterogeneity is qualitatively consistent with our simulation results, firm-level estimates of acceleration time vary substantially between the two approaches.

Table 3. Regression Analysis of Sales-Weighted Average Acceleration Performance over the Period 1976–2014[a]

	(1)	(2)	(3)	(4)
time trend	−0.181[b]	−0.174[b]	−0.182[b]	−0.176[b]
	(0.014)	(0.012)	(0.014)	(0.012)
D1	0.523[b]	0.161		
	(0.202)	(0.165)		
D2	0.571[b]	0.467[b]		
	(0.180)	(0.166)		
real gasoline prices		0.261[d]		0.270[d]
		(0.158)		(0.148)
B1			0.195[b]	0.070
			(0.070)	(0.049)
B2			−0.074	0.082
			(0.061)	(0.053)
Constant	15.008[b]	14.293[b]	15.024[b]	14.307[b]
	(0.363)	(0.553)	(0.362)	(0.514)
imputed impact of the policy on acceleration (s)	1.1	0.63	1.1	0.71
joint F-test	5.37[c]	4.01[c]	8.22[b]	8.57[c]
R^2	0.750	0.754	0.753	0.757
number of observations	493	493	493	493

[a]The unit of observation is a firm-year-vehicle type. [b]$p < 0.01$. [c]$p < 0.05$. [d]$p < 0.1$.

In sum, the trajectories in acceleration performance we observe are qualitatively consistent with our simulation results; following the introduction of the reformed policy, observed acceleration performance is significantly worse than our counterfactual estimate based on trends before the policy change. Our estimated impact of the reformed standards on sales-weighted average acceleration performance are very similar across our econometric and simulation results, although the firm-level results are not as congruent. These results lend further support to our hypothesis that acceleration trade-offs play an important role in automakers' compliance strategies.

CONCLUSION

Environmental policies can significantly influence engineering design decisions as firms reoptimize their products to meet compliance requirements at minimum cost. We evaluate the potential importance of vehicle design trade-offs between fuel economy and acceleration performance in automakers' responses to the reformed CAFE and GHG standards. Using simulations of the automotive industry, we find that automakers have an incentive to use these design trade-offs and that GHG emissions and compliance costs (measured in terms of lost producer profits and consumer surplus) are significantly lower when these trade-offs are accounted for. We also find that these simulation-based estimates are consistent with changes in vehicle attributes observed in the years following the announcement of the policy. Given the potential importance of acceleration trade-offs as a means of complying with vehicle standards, regulatory agencies should consider these performance trade-offs. Our results also imply that previous analyses of the regulations that do not include these trade-offs may

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307–10315

significantly overestimate compliance costs and underestimate GHG emission reductions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.7b03743.

Detailed information on each of the constituent data and models used in our analysis as well as descriptions of several robustness checks on our results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kwhitefoot@cmu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful for financial support from Carnegie Mellon University and from valuable comments from Jeremy Michalek, Jim Sallee, and seminar participants at Carnegie Mellon University, University of California-Berkeley, and the University of Chicago.

■ REFERENCES

(1) Average fuel economy standards. *U.S.C* **1994**, *49*, §32902.

(2) U.S. Environmental Protection Agency; U.S. Department of Transportation National Highway Traffic Safety Administration. *Joint Technical Support Document: Final Rulemaking for 2017–2025 Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards*; EPA-420-R-12-901; 2012.

(3) U.S. Department of Transportation National Highway Traffic Safety Administration. *Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012-MY 2016 Passenger Cars and Light Trucks*; Office of Regulatory Analysis and Evaluation National Center for Statistics and Analysis; 2010; p 47.

(4) Cheah, L. W.; Bandivadekar, A. P.; Bodek, K. M.; Kasseris, E. P.; Heywood, J. B. The trade-off between automobile acceleration performance, weight, and fuel consumption. *SAE International Journal of Fuels and Lubricants* **2009**, *1* (1), 771–777.

(5) Knittel, C. R. Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector. *American Economic Review* **2011**, *2012* (101), 3368–3399.

(6) Klier, T.; Linn, J. The effect of vehicle fuel economy standards on technology adoption. *Journal of Public Economics* **2016**, *133*, 41–63.

(7) Jenn, A.; Azevedo, I. M. L.; Michalek, J. J. Alternative Fuel Vehicle Adoption Increases Fleet Gasoline Consumption and Greenhouse Gas Emissions under United States Corporate Average Fuel Economy Policy and Greenhouse Gas Emissions Standards. *Environ. Sci. Technol.* **2016**, *50* (5), 2165–2174.

(8) Anderson, S. T.; Sallee, J. M. Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards. *American Economic Review* **2011**, *101* (4), 1375–1409.

(9) Kleit, A. N. The effect of annual changes in automobile fuel economy standards. *Journal of Regulatory Economics* **1990**, *2* (2), 151–172.

(10) Goldberg, P. K. Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry. *Econometrica* **1995**, *63* (4), 891.

(11) Austin, D.; Dinan, T. Clearing the air: The costs and consequences of higher CAFE standards and increased gasoline taxes. *Journal of Environmental Economics and Management* **2005**, *50* (3), 562–582.

(12) Whitefoot, K. S.; Skerlos, S. J. Design incentives to increase vehicle size created from the US footprint-based fuel economy standards. *Energy Policy* **2012**, *41*, 402–411.

(13) U.S. Environmental Protection Agency; U.S. Department of Transportation National Highway Traffic Safety Administration. Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule. *Federal Register* 2010, *75* (88).

(14) US EPA, O. Clean Air Act Vehicle and Engine Enforcement Case Resolutions https://www.epa.gov/enforcement/clean-air-act-vehicle-and-engine-enforcement-case-resolutions (accessed May 27, 2016).

(15) Shiau, C.-S. N.; Michalek, J. J.; Hendrickson, C. T. A structural analysis of vehicle design responses to Corporate Average Fuel Economy policy. *Transportation Research Part A: Policy and Practice* **2009**, *43* (9), 814–828.

(16) Ito, K.; Sallee, J. M. *The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards*; National Bureau of Economic Research, 2014.

(17) Klier, T.; Linn, J. New-vehicle characteristics and the cost of the Corporate Average Fuel Economy standard. *RAND Journal of Economics* **2012**, *43* (1), 186–213.

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

(19) Nelson, S. A.; Parkinson, M. B.; Papalambros, P. Y. Multicriteria Optimization in Product Platform Design. *Journal of Mechanical Design* **2001**, *123* (2), 199.

(20) Kim, H. M.; Kokkolaras, M.; Louca, L. S.; Delagrammatikas, G. J.; Michelena, N. F.; Filipi, Z. S.; Papalambros, P. Y.; Stein, J. L.; Assanis, D. N. Target cascading in vehicle redesign: a class VI truck study. *International Journal of Vehicle Design* **2002**, *29* (3), 199–225.

(21) Orsborn, S.; Cagan, J.; Boatwright, P. Quantifying Aesthetic Form Preference in a Utility Function. *Journal of Mechanical Design* **2009**, *131* (6), 61001.

(22) Michalek, J. J.; Feinberg, F. M.; Papalambros, P. Y. Linking marketing and engineering product design decisions via analytical target cascading. *Journal of Product Innovation Management* **2005**, *22* (1), 42–62.

(23) Zhao, Y.; Thurston, D. Maximizing Profits From End-of-Life and Initial Sales With Heterogeneous Consumer Demand. *Journal of Mechanical Design* **2013**, *135* (4), 41001.

(24) Kwak, M.; Kim, H. Market positioning of remanufactured products with optimal planning for part upgrades. *Journal of Mechanical Design* **2013**, *135* (1), 11007.

(25) Shiau, C.-S. N.; Michalek, J. J. Optimal Product Design Under Price Competition. *Journal of Mechanical Design* **2009**, *131* (7), 71003.

(26) Morrow, W. R.; Mineroff, J.; Whitefoot, K. S. Numerically Stable Design Optimization With Price Competition. *Journal of Mechanical Design* **2014**, *136* (8), 81002.

(27) Michalek, J. J.; Papalambros, P. Y.; Skerlos, S. J. A Study of Fuel Efficiency and Emission Policy Impact on Optimal Vehicle Design Decisions. *Journal of Mechanical Design* **2004**, *126* (6), 1062.

(28) Frischknecht, B. D.; Whitefoot, K. S.; Papalambros, P. Y. On the Suitability of Econometric Demand Models in Design for Market Systems. *Journal of Mechanical Design* **2010**, *132* (12), 121007.

(29) Shiau, C.-S. N.; Michalek, J. J. Should Designers Worry About Market Systems? *Journal of Mechanical Design* **2009**, *131* (1), 11011.

(30) Braess, H.-H.; Seiffert, U. *Handbook of Automotive Engineering*; SAE, 2005.

(31) Sörensen, D. *The Automotive Development Process*; Springer, 2006.

(32) Weber, J. *Automotive development processes*; Springer, 2014.

(33) NHTSA. *Preliminary Regulatory Impact Analysis: Corporate Average Fuel Economy for MY2011–2015 Passenger Cars and Light Trucks*; U.S. Department of Transportation, 2008.

(34) EPA; DOT NHTSA. 2017 and Later Model Year Light-Duty Vehicle Greenhouse Gas Emissions and Corporate Average Fuel Economy Standards. *Federal Register* **2012**, *77* (199), 62712.

(35) Berry, S. T. Estimating discrete-choice models of product differentiation. *RAND Journal of Economics* **1994**, *25*, 242–262.
(36) Berry, S.; Levinsohn, J.; Pakes, A. Automobile prices in market equilibrium. *Econometrica* **1995**, *63*, 841–890.
(37) Berry, S.; Levinsohn, J.; Pakes, A. Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy* **2004**, *112* (1), 68–105.
(38) Train, K. E.; Winston, C. Vehicle choice behavior and the declining market share of us automakers*. *International Economic Review* **2007**, *48* (4), 1469–1496.
(39) Jacobsen, M. R. Evaluating US fuel economy standards in a model with producer and household heterogeneity. *American Economic Journal: Economic Policy* **2013**, *5* (2), 148–187.
(40) Small, K. A.; Van Dender, K. Fuel efficiency and motor vehicle travel: the declining rebound effect. *Energy Journal* **2007**, *28*, 25–51.
(41) U.S. Department of Transportation National Highway Traffic Safety Administration. *Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012-MY 2016 Passenger Cars and Light Trucks*, 2010.
(42) US Environmental Protection Agency. *Light-Duty Automotive Technology, Carbon Dioxide Emissions, and Fuel Economy Trends: 1975 through 2015*; EPA-420-R-15-016, 2015.
(43) U.S. Environmental Protection Agency. U.S. Department of Transportation National Highway Traffic Safety Administration. Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule. *Federal Register* **2010**, *75* (88), 25413–25414.

DOI: 10.1021/acs.est.7b03743
Environ. Sci. Technol. 2017, 51, 10307–10315

Transportation Research Part D 65 (2018) 194–212

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Re-searching for hidden costs: Evidence from the adoption of fuel-saving technologies in light-duty vehicles

Hsing-Hsiang Huang[a,*], Gloria Helfand[b], Kevin Bolon[b], Robert Beach[c], Mandy Sha[c], Amanda Smith[c]

[a] *Oak Ridge Institute for Science and Education, at the Office of Transportation and Air Quality, U.S. Environmental Protection Agency, 2000 Traverwood Dr, Ann Arbor, MI 48105, USA*
[b] *Assessment and Standards Division, Office of Transportation and Air Quality, U.S. Environmental Protection Agency, 2000 Traverwood Dr, Ann Arbor, MI 48105, USA*
[c] *RTI International, 3040 E. Cornwallis Rd, Research Triangle Park, NC 27709, USA*

ARTICLE INFO

Keywords:
Light-duty vehicles
Fuel economy standards
Vehicle greenhouse gas standards
Content analysis
Vehicle fuel-saving technology
Vehicle operational characteristics

ABSTRACT

A variety of fuel-saving technologies have been implemented in light-duty vehicles since 2012 under the U.S. Environmental Protection Agency's (EPA) and Department of Transportation (DOT)'s light-duty vehicle greenhouse gas emissions and fuel economy standards. Questions have arisen whether there are hidden costs that have not been included in the net benefit calculations as a result of adoption of the new technologies. In this paper, we replicate and expand results from Helfand et al. (2016). We define hidden costs of the new technologies as problems with operational characteristics such as acceleration, handling, ride comfort, noise, braking feel, and vibration, not all of which are easily measured by objective criteria. We overcome the empirical challenge by using data coded from online professional auto reviews that qualitatively evaluate fuel-saving technologies and operational characteristics for model years 2014 and 2015 vehicles. We estimate relationships of fuel-saving technologies and operational characteristics, including an overall vehicle assessment, and find little correlation of hidden costs with the technologies themselves. Variable quality of implementation of technologies across automakers may better explain negatively evaluated operational characteristics. The results imply that automakers have typically been able to implement fuel-saving technologies without harm to vehicle operational characteristics.

1. Introduction

Fossil fuel combustion in transportation has contributed approximately one-fourth of greenhouse gas (GHG) emissions in the United States in recent years (U.S. Environmental Protection Agency (EPA), 2016a). In an effort to reduce GHG emissions and improve energy security, the U.S. Environmental Protection Agency (EPA) and the Department of Transportation (DOT) established vehicle GHG and fuel economy standards for light-duty vehicles for model years (MYs) 2012 through 2025. In the presence of the standards, vehicle manufacturers have implemented a wide range of fuel-saving technologies (EPA, 2016b; EPA, DOT and California

* Corresponding author.
E-mail addresses: hsihuang@umich.edu (H.-H. Huang), helfand.gloria@epa.gov (G. Helfand), bolon.kevin@epa.gov (K. Bolon), rbeach@rti.org (R. Beach), msha@rti.org (M. Sha), acsmith@rti.org (A. Smith).

https://doi.org/10.1016/j.trd.2018.08.009

Air Resources Board (CARB) 2016, Chapter 5).[1] Assessments of the standards have found enormous net benefits to society, including significant net benefits from fuel savings for new vehicle buyers (e.g., EPA and DOT, 2010, 2012; EPA, DOT, and CARB, 2016). However, questions have been raised about whether there are hidden costs that have not been included in the net benefit calculations as a result of adoption of the new technologies (Allcott and Greenstone, 2012; Gillingham and Palmer, 2014; Helfand et al., 2016). In particular, hidden costs that exceed the net positive financial benefits from fuel reduction for new vehicle buyers might explain why markets had been slow to adopt fuel-saving technologies on light-duty vehicles in the absence of the standards. Given the wide range of fuel-saving technologies developed and adopted in recent years, it is important to understand whether any of the new technologies impose hidden costs.

We consider hidden costs to be negative impacts of the technologies on performance, drivability, ride comfort, and other characteristics that would cause losses to consumer welfare and are difficult to measure. For instance, if six-speed automatic transmissions were clunky or otherwise worse to drive than traditional four-speed automatic transmissions, buyers of vehicles with six-speed transmissions would suffer welfare losses from the hidden costs and thus would be less interested in buying them. As the effects of the GHG standards depend critically on consumers buying vehicles with fuel-saving technologies, an evaluation of the new technologies should consider potential hidden costs.

One set of literature relevant to hidden costs as a result of adoption of fuel-saving technology has focused on estimating the tradeoffs between fuel economy and horsepower and weight (e.g., Knittel, 2011; Klier and Linn, 2012, 2016; MacKenzie and Heywood, 2015). This literature has focused on estimating this relationship as technological, not involving consumer response; see EPA, DOT, and CARB (2016), Chapter 4.1.3, for further discussion.

This paper is closely related to Helfand et al. (2016), which investigated whether there are hidden costs in a range of operational characteristics. Though operational characteristics that consumers may care about are not well measured by quantified vehicle attributes, they are usually evaluated qualitatively by professional auto reviewers. Helfand et al. (2016) gathered data on both operational characteristics and fuel-saving technologies for MY 2014 by conducting a content analysis of online auto reviews of MY 2014 vehicles. Content analysis involves systematic coding of text; it can be used to convert qualitative information to quantitative (Krippendorff, 2013). They did not find systematic evidence of negative operational characteristics associated with adoption of a variety of fuel-saving technologies, suggesting that it is possible to use the technologies on light-duty vehicles without imposing hidden costs on consumers.

This paper builds on Helfand et al. (2016) in several ways. First, this paper adds evaluations from professional auto reviews for MY 2015 vehicles to the dataset. These additional data provide an opportunity for validation of the results of the original study.

Second, instead of only using cross-sectional variation in technology adoption, the use of year fixed effects allows for control of more unobserved factors that may be correlated with changes in technology adoption and evaluation results, such as changes in consumer preferences. The larger dataset also helps avoid small sample size for some technologies. The estimation results of this paper, using the pooled data and adjusted standard errors dealing with potential small sample bias, are consistent with Helfand et al. (2016)'s conclusion that fuel-saving technologies generally appear not to be associated with negative operational impacts.

Third, to further explore the role of variable implementation quality for fuel-saving technologies, proposed by Helfand et al. (2016), we estimate whether negative evaluations of operational characteristics are correlated with negatively reviewed technologies, conditional on the presence of the technologies. We find evidence of positive relationships between negatively evaluated technologies and negatively evaluated operational characteristics, suggesting that poorly implemented technologies, instead of the presence of the technologies themselves, may be correlated with hidden costs.

Lastly, we examine whether fuel-saving technologies are associated with the overall assessment of the reviewed vehicles concluded by each reviewer. An overall rating, advising whether to purchase the vehicle, may be explicit, or it may be inferred from the evaluation of vehicle characteristics and comparison with vehicles in the same segment. The overall rating is expected to include any factors that the reviewer may consider, even if they are not specifically evaluated. We do not find evidence of associations between negative overall ratings of vehicles and the presence of fuel-saving technologies. Instead, we find negatively rated technologies and negatively rated operational characteristics are highly associated with the overall rating, further suggesting that lower quality technologies, with their adverse effects on operational characteristics, play key roles in getting a negative overall assessment.

These results suggest that, to date, automakers generally have been able to implement fuel-saving technologies without imposing hidden costs on consumers. This finding implies that net benefits from fuel savings suggested in the literature are higher than potential hidden costs of adoption of fuel-saving technology.

The remainder of this paper is structured as follows. The next section describes our data. Section 3 covers our estimation approach. Section 4 presents our results for the relationship of fuel-saving technologies and operational impacts. Section 5 presents our results for the relationship of the technologies with the overall assessment. Section 6 concludes.

2. Data and content analysis

The data for this study come from online professional auto reviews of MY 2014 and MY 2015 new vehicles. Professional auto reviews provide qualitative evaluation of both technologies and operational characteristics. For both of these categories, quality may be difficult to quantify but is very important to consumers. Hidden costs emerge if negative impacts on these operational

[1] For instance, gasoline direct injection was widely used on nearly half of all vehicles in model year 2016, while it was used in less than 3% of vehicles in model year 2008 (EPA, 2016b).

characteristics exist as a result of adoption of fuel-saving technology.

Content analysis provides a systematic approach to evaluate reviewers' evaluations of the quality of fuel-saving technologies and operational characteristics of the vehicles they review. This method involves breaking text into words and phrases that can be categorized and analyzed using specified definitional codes (Krippendorff 2013). Content analysis has been widely used in the humanities and social sciences to classify, measure, and evaluate themes and symbols in various communications media. See Sha and Beach (2015) and Sha et al. (2016) for further background and detail, and Helfand et al. (2016) for further examples of vehicle-related content analyses.

2.1. Identification of relevant websites of professional auto reviews

As detailed in Sha and Beach (2015) and Sha et al. (2016), we followed a set of specific procedures to identify the websites used in this study. In particular, we aimed for websites that contained reviews from professional auto reviewers and that consumers are most likely to consult when making vehicle buying decisions in the United States. First, using Google and Yahoo internet search engines, we sought websites on the first page of search returns for keywords "new cars," "buying a new car," and "auto reviews." Second, we excluded websites that did not have national and professional auto reviews. Third, we used monthly unique views from Quantcast.com and Compete.com to gauge website popularity, and excluded websites that had less than one million unique views in both Quantcast.com and Compete.com. Finally, we screened websites to include only professional reviews that evaluated vehicles and technologies. Each review must have gone beyond a basic specification list, have an independent assessment of vehicle quality, and show evidence of the reviewer having test-driven the reviewed vehicle.

For MY 2014 vehicles, six websites were selected by following the sampling procedures above: Automobile Magazine, Auto Trader, Car and Driver, Consumer Reports, Edmunds, and Motor Trend. For MY 2015 vehicles, we started with the six websites for MY 2014 vehicles, and followed the same procedures to identify other potential websites. One new website, Cars.com, was added in MY 2015, because its web viewership met our criteria for inclusion. As in Helfand et al. (2016), this study included all reviews of new MY 2014 and 2015 vehicles subject to the light-duty GHG standards. We dropped the reviews of Volkswagen and Audi diesel vehicles due to concerns over compliance with emissions standards, as well as medium-duty vehicles not subject to the light-duty vehicle standards. Table 1 reports the number of reviews by website in our analysis. Our dataset includes 2238 separate reviews over the two model years, including 1003 for MY 2014 and 1235 reviews for MY 2015.

The vehicles reviewed in our sample appear to be roughly representative of vehicles offered, based on data from fueleconomy.gov (U.S. Department of Energy and EPA, 2014, 2015), although the reviews do not reflect sales (see Table 2). Vehicles offered may be a better comparison group than sales, because potential buyers examine auto reviews based on the choice set, rather than what other people buy. Grouped at the vehicle class level as presented in Table 3, the percentage of auto reviews by class is roughly similar to the national fleet-wide breakdown (again based on fueleconomy.gov data) of MY 2014 and MY 2015 vehicles. While the number of reviews of mid-sized cars are over-represented for MY 2014 vehicles, it becomes slightly under-represented for MY 2015 vehicles.

2.2. Coding qualitative assessments of vehicle characteristics

This study codes both fuel-saving technologies and operational characteristics discussed in auto reviews. The set of technologies coded included most of the technologies proposed for compliance purposes in EPA and DOT (2010, 2012). The set of operational characteristics was developed from judgment of factors likely to be relevant to drivers, with refinements based on experience with the reviews. To ensure consistency of coding between MY 2014 and MY 2015 vehicles, and thus allow for better assessment of replication of results, the same coders and coding definitions of fuel-saving technologies and operational characteristics were used for both samples (Sha et al. 2016). One new fuel-saving technology, fuel cell, was added for MY 2015. Tables 4 and 5 list the coded fuel-saving technologies and operational characteristics, respectively, in the study. A hybrid vehicle is a special case, because all hybrids have stop-start and CVT. For this study, stop-start and CVT are only possible for non-hybrid vehicles; "hybrid" is considered a package including stop-start and CVT.

Coding processes were also consistent over both datasets. For each auto review, every mentioned fuel-saving technology and

Table 1
Auto reviews by website.

Website	MY 2014		MY 2015		Pooled	
	Review Count	%	Review Count	%	Review count	%
automobilemag.com	144	14	138	11	282	13
autotrader.com	224	22	336	27	560	25
caranddriver.com	216	22	202	16	418	19
cars.com	0	0	90	7	90	4
consumerreports.org	86	9	79	6	165	7
edmunds.com	112	11	105	9	217	10
motortrend.com	221	22	285	23	506	23
Total	1003	100	1235	100	2238	100

Table 2
Auto reviews by make, compared with fueleconomy.gov counts.

Make	MY 2014				MY 2015			
	Auto Review		fueleconomy.gov		Auto Review		fueleconomy.gov	
	Count	%	Count	%	Count	%	Count	%
Acura	24	2.4	16	1.3	22	1.8	10	0.8
Audi	37	3.7	48	3.9	60	4.9	55	4.3
BMW	69	6.9	98	8.0	77	6.2	121	9.4
Bentley	11	1.1	7	0.6	16	1.3	8	0.6
Buick	27	2.7	16	1.3	11	0.9	16	1.3
Cadillac	36	3.6	35	2.8	21	1.7	29	2.3
Chevrolet	85	8.5	77	6.3	101	8.2	92	7.1
Chrysler	4	0.4	14	1.1	28	2.3	13	1.0
Dodge	24	2.4	35	2.8	41	3.3	35	2.7
Ferrari	7	0.7	13	1.1	0	0.0	14	1.1
Fiat	8	0.8	7	0.6	4	0.3	10	0.8
Ford	47	4.7	88	7.2	79	6.4	78	6.1
GMC	17	1.7	36	2.9	21	1.7	51	4.0
Honda	34	3.4	30	2.4	30	2.4	27	2.1
Hyundai	19	1.9	38	3.1	64	5.2	44	3.4
Infiniti	25	2.5	29	2.4	23	1.9	31	2.4
Jaguar	28	2.8	20	1.6	22	1.8	23	1.8
Jeep	42	4.2	35	2.8	15	1.2	39	3.0
Kia	44	4.4	35	2.8	44	3.6	38	2.9
Lamborghini	0	0.0	7	0.6	5	0.4	4	0.3
Land Rover	15	1.5	13	1.1	17	1.4	11	0.9
Lexus	23	2.3	25	2.0	54	4.4	32	2.5
Lincoln	6	0.6	16	1.3	22	1.8	21	1.6
Maserati	0	0.0	6	0.5	1	0.1	6	0.5
Mazda	49	4.9	25	2.0	15	1.2	24	1.9
Mercedes-Benz	74	7.4	85	6.9	84	6.8	83	6.5
Mini Cooper	11	1.1	46	3.7	9	0.7	44	3.4
Mitsubishi	17	1.7	19	1.5	10	0.8	18	1.4
Nissan	40	4.0	51	4.1	54	4.4	50	3.9
Porsche	34	3.4	52	4.2	47	3.8	61	4.8
Ram	7	0.7	13	1.1	8	0.6	15	1.2
Rolls Royce	9	0.9	7	0.6	4	0.3	7	0.6
Scion	4	0.4	9	0.7	8	0.6	7	0.6
Smart	1	0.1	4	0.3	0	0.0	0	0.0
Subaru	25	2.5	23	1.9	59	4.8	21	1.6
Tesla	0	0.0	3	0.2	4	0.3	8	0.6
Toyota	63	6.3	58	4.7	75	6.1	53	4.1
Volkswagen	32	3.2	50	4.1	44	3.6	46	3.6
Volvo	5	0.5	13	1.1	36	2.9	21	1.6
Other*	0	0.0	27	2.2	0	0.0	18	1.4
Total	1003		1229		1235		1284	

* Other includes Alfa Romeo, Aston Martin, Bugatti, BYD, Lotus, McLaren, Mobility Ventures LLC, Pagani, Roush, and SRT.

operational characteristic was coded as Positive, Negative, or Neutral. For instance, a passage of text containing a negative evaluation of stop-start technology would be coded as Negative – Stop-Start, while another passage of text containing a positive evaluation of steering feel was coded as Positive – Steering Feel. It was coded as Neutral when the reviewer did not demonstrate an intensity of opinion that could be clearly discerned to be positive or negative.

In addition to categorizing each mention of a technology and operational characteristic as positive, negative or neutral, the overall assessment, or recommendation, of the review was coded as well. Review summaries and conclusions often provide a general idea of the reviewer's attitude about the reviewed vehicle (e.g. a reviewer might opine whether consumers should/shouldn't purchase a particular car.) Reading the review as a whole led to assigning a positive, negative, or mixed evaluation to each review to capture the overall assessment.

The analysis here uses what Helfand et al. (2016) call review-level data: that is, the codes are aggregated for each review. If a specific technology or characteristic is mentioned multiple times in an auto review and all the codes for the mentions of the technology are the same throughout the auto review, then it is listed once in the relevant column. For instance, if all of the mentions of turbocharging are negative in an auto review, then that review is coded once for Turbocharged – Negative. On the other hand, if a technology or characteristic receives more than one kind of evaluation – for instance, both positive and negative evaluations in the same auto review – the review is coded once for each evaluation of the technology – in this case, once positive and once negative. Helfand et al. (2016) found little difference for results when using individual codes compared to the review-level analysis.

Table 3
Auto reviews by vehicle class, compared with fueleconomy.gov counts.

Vehicle Class	MY 2014				MY 2015			
	Auto Review		fueleconomy.gov		Auto Review		fueleconomy.gov	
	Count	%	Count	%	Count	%	Count	%
Subcompact Cars	84	8.4	101	8.2	111	9.0	107	8.3
Minicompact Cars	11	1.1	52	4.2	13	1.1	56	4.4
Compact Cars	181	18.1	201	16.4	271	21.9	207	16.1
Two Seaters	88	8.8	93	7.6	75	6.1	100	7.8
Midsize Cars	227	22.6	215	17.5	149	12.1	203	15.8
Large Cars	93	9.3	104	8.5	119	9.6	113	8.8
Small Station Wagons	26	2.6	36	2.9	56	4.5	35	2.7
Midsize Station Wagons	6	0.6	4	0.3	21	1.7	8	0.6
Passenger Vans	1	0.1	16	1.3	1	0.1	14	1.1
Minivans	15	1.5	14	1.1	28	2.3	14	1.1
Small SUVs	130	13.0	179	14.6	212	17.2	196	15.3
Standard SUVs	94	9.4	116	9.4	121	9.8	134	10.4
Small Pickup Trucks	1	0.1	14	1.1	27	2.2	24	1.9
Standard Pickup Trucks	42	4.2	54	4.4	28	2.3	45	3.5
Other*	4	0.4	30	2.4	3	0.2	28	2.2
Total	1003	100	1229	100	1235	100	1284	100

* Other includes special purpose vehicle and cargo vans.

Table 4
Total number of positive, negative, and neutral evaluations of fuel-saving technologies by auto review.

Fuel-Saving Technology	MY 2014							MY 2015							Pooled
	Negative		Neutral		Positive		Total	Negative		Neutral		Positive		Total	Total
Active Air Dam	0	0%	0	0%	6	100%	6	0	–	0	–	0	–	0	6
Active Grill Shutters	0	0%	0	0%	1	100%	1	1	14%	0	0%	6	86%	7	8
Active Ride Height	0	0%	1	33%	2	67%	3	0	–	0	–	0	–	0	3
Low Resistance Tires	4	24%	5	29%	8	47%	17	4	31%	1	8%	8	62%	13	30
Electronic Power Steering	45	22%	42	20%	121	58%	208	22	14%	19	12%	116	74%	157	365
Turbocharged	20	9%	23	10%	180	81%	223	43	13%	35	10%	264	77%	342	565
GDI	6	9%	6	9%	54	82%	66	4	6%	6	9%	55	85%	65	131
Cylinder Deactivation	1	3%	4	11%	30	86%	35	4	16%	3	12%	18	72%	25	60
Diesel	7	12%	9	15%	44	73%	60	5	28%	2	11%	11	61%	18	78
Hybrid	16	23%	10	14%	45	63%	71	10	21%	5	11%	32	68%	47	118
Plug-In Hybrid Electric	4	14%	6	21%	18	64%	28	4	22%	3	17%	11	61%	18	46
Full Electric	2	9%	6	27%	14	64%	22	0	0%	3	15%	17	85%	20	42
Fuel Cell	0	–	0	–	0	–	0	0	0%	0	0%	1	100%	1	1
Stop-Start	14	27%	7	14%	30	59%	51	15	31%	9	19%	24	50%	48	99
High Speed Automatic	60	14%	81	20%	273	66%	414	96	20%	76	16%	310	64%	482	896
CVT	35	31%	20	18%	57	51%	112	38	30%	14	11%	75	59%	127	239
DCT	16	24%	10	15%	42	62%	68	18	17%	10	10%	77	73%	105	173
Elec Assist/Low Drag Brakes	1	14%	3	43%	3	43%	7	0	0%	0	0%	2	100%	2	9
Lighting-LED	1	5%	2	10%	17	85%	20	0	0%	1	4%	25	96%	26	46
Mass Reduction	0	0%	9	12%	65	88%	74	3	6%	2	4%	43	90%	48	122
Passive Aerodynamics	4	10%	7	18%	29	73%	40	2	11%	0	0%	17	89%	19	59
Efficiency Totals	378	16%	391	16%	1668	68%	2437	463	18%	310	12%	1783	70%	2556	4993

2.3. Summary statistics

Table 4 reports the number of auto reviews that have positive, negative, or neutral evaluations of fuel-saving technologies. The fuel technologies examined are not mentioned very frequently in the reviews. There are about 1.52 and 1.27 codes of fuel-saving technologies per review for the MY 2014 and MY 2015 data, respectively. The most mentioned technologies are high speed automatic, turbocharged, electronic power steering, and continuously variable transmissions (CVT). Among the four most evaluated technologies, turbocharged and high speed automatic have substantially more mentions for MY 2015 vehicles than MY 2014 vehicles.

In the data, positive evaluations exceed negative evaluations for all the technologies examined for both years. As reported in Table 4, in the aggregate, positive evaluations are about 70% of the totals, while negative evaluations are less than 20%. CVT, stop-start, and low rolling resistance tires are the most frequently negatively reviewed fuel-saving technologies for the data. However, even these most frequently negatively reviewed technologies have majority positive evaluations. For example, CVT has 51% and 59% positive evaluations for the MY 2014 and MY 2015 data, respectively, while it has about 30% negative evaluations for vehicles of

Table 5
Total number of positive, negative, and neutral evaluations of operational characteristics by auto review.

Operational Characteristics	MY 2014							MY 2015							Pooled
	Negative		Neutral		Positive		Total	Negative		Neutral		Positive		Total	Total
Handling															
Steering Feel	147	20%	163	22%	442	59%	752	173	21%	119	15%	517	64%	809	1561
Cornering Ability	92	14%	116	17%	471	69%	679	135	18%	131	17%	500	65%	766	1445
General Drivability	116	15%	146	18%	531	67%	793	173	19%	130	14%	621	67%	924	1717
General Handling	82	12%	130	20%	450	68%	662	139	17%	116	14%	571	69%	826	1488
Acceleration															
Acceleration Feel	76	15%	73	15%	343	70%	492	158	26%	47	8%	405	66%	610	1102
Acceleration Capability	164	16%	231	23%	630	61%	1025	254	21%	232	19%	744	60%	1230	2255
General Acceleration	24	17%	27	19%	89	64%	140	47	18%	39	15%	170	66%	256	396
Braking															
Brake Feel	46	13%	58	17%	246	70%	350	99	27%	46	12%	226	61%	371	721
Stopping Ability	31	9%	78	22%	249	70%	358	49	14%	73	21%	228	65%	350	708
General Braking	21	17%	18	15%	83	68%	122	17	17%	14	14%	72	70%	103	225
Noise															
Tire-Road Noise	72	24%	74	25%	153	51%	299	113	34%	48	14%	172	52%	333	632
Wind Noise	29	14%	46	21%	139	65%	214	50	21%	36	15%	147	63%	233	447
Interior Noise	16	32%	6	12%	28	56%	50	20	40%	4	8%	26	52%	50	100
Powertrain Noise	145	25%	104	18%	330	57%	579	149	24%	68	11%	400	65%	617	1196
General Noise	58	14%	33	8%	332	78%	423	61	15%	24	6%	332	80%	417	840
Vibration															
Chassis Vibration	7	70%	3	30%	0	0%	10	2	67%	0	0%	1	33%	3	13
Powertrain Vibration	10	40%	8	32%	7	28%	25	10	56%	2	11%	6	33%	18	43
General Vibration	19	42%	13	29%	13	29%	45	14	42%	2	6%	17	52%	33	78
Ride Comfort	149	19%	171	22%	462	59%	782	201	22%	135	15%	585	64%	921	1703
Fuel Economy	161	22%	178	24%	396	54%	735	239	27%	163	19%	474	54%	876	1611
Range	7	16%	11	24%	27	60%	45	7	18%	8	21%	23	61%	38	83
Charging	3	30%	0	0%	7	70%	10	2	14%	2	14%	10	71%	14	24
Operational Totals	1475	17%	1687	20%	5428	63%	8590	2112	22%	1439	15%	6247	64%	9798	18388

both model years. These results suggest that it is possible to implement these technologies without significant hidden costs.

Some technologies have limited reviews in a single model year data, such as plug-in hybrid electric, passive aerodynamics, and low resistance tires. As shown in the last column of Table 4, all but five technologies have more than 30 reviews in the pooled data; the exceptions are active air dam, active grill shutters, active ride height, fuel cell, and electric assist or low drag brakes. The technologies with greater than 30 reviews most frequently mentioned positively in percentage terms include LED lighting, mass reduction, gasoline direct injection (GDI), cylinder deactivation, turbocharged, and passive aerodynamics. The technologies with greater than 30 reviews least frequently mentioned negatively in percentage terms are the same ones, except that full electric replaces turbocharged. The most frequently negatively reviewed technologies over the two years by percentage are CVT, stop-start, low-rolling-resistance tires, hybrid, and dual-clutch transmissions (DCT). As noted, though, these all are rated positively for more than 50% of the reviews where they are mentioned.

As reported in Table 5, mentions of the operational characteristics in the aggregate have more than 60% positive evaluations, about 20% neutral evaluations, and about 20% negative evaluations. The reviews of operational characteristics are slightly more negative (17–22% negative) than the reviews of fuel-saving technologies (16–18% negative). Among the operational characteristics, chassis, powertrain, and general vibration have the highest percentage of negative reviews across both model years, followed by interior and tire-road noise. Mentions of vibration are relatively infrequent; only charging for plug-in electric vehicles is mentioned as infrequently. It may be that vibration is mentioned only when there is a problem.

Fig. 1 shows a summary of the overall assessments of the vehicles reviewed. Similar to the aggregated operational characteristics, about 65% of vehicles are positively reviewed on the overall assessment. While MY 2015 vehicles have slightly more mixed evaluations than MY 2014 vehicles, only about 8% of the reviews have an overall negative evaluation.

In Panel (A) of Fig. 2, we divide our pooled data into two groups: one (red) includes the auto reviews that mention the technology listed on the vertical axis, and the other (blue) includes the auto reviews that do not mention the technology. Then we compare the shares of auto reviews with an overall negative assessment between the two groups. For instance, while 8% of reviews that do not mention GDI have a negative overall assessment, only 2% of reviews that do mention GDI have a negative overall assessment. As Panel A indicates, vehicles with most fuel-saving technologies are less likely to have negative overall assessments than vehicles without the technologies, with the exception of vehicles with CVT, hybrid, low-rolling-resistance tires, LED lights, and high speed automatic.

In Panel (B) of Fig. 2, we compare the total number of negative evaluations of operational characteristics between the two groups. It suggests that vehicles with most fuel-saving technologies get fewer negative evaluations of operational characteristics than vehicles

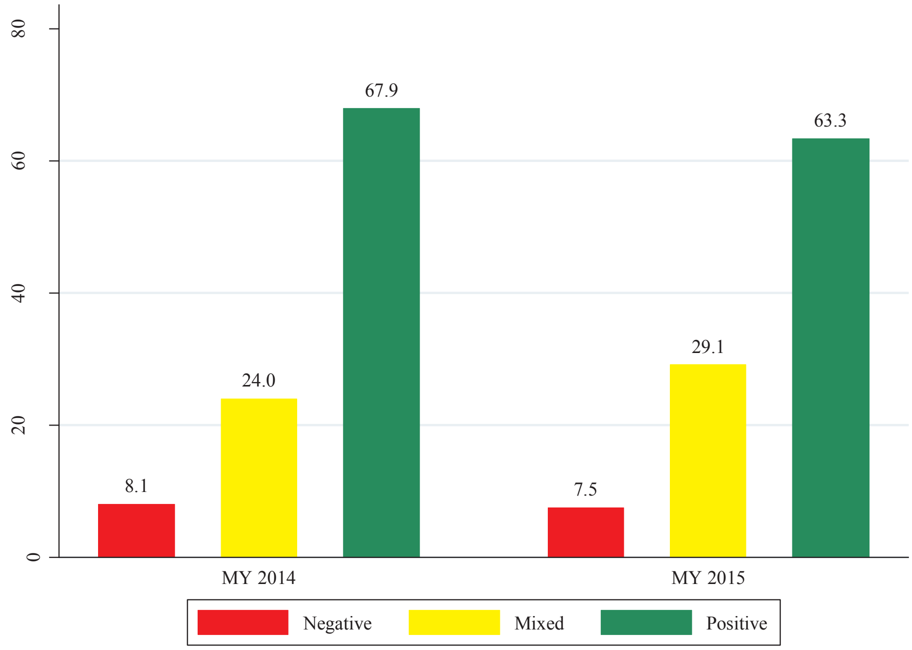


Fig. 1. Overall assessment of the quality of the vehicle reviewed.

without the technologies, with the exception of CVT, hybrid, low-rolling-resistance tires, and plug-in hybrid-electric vehicles.

It is important to note that the difference between the two groups does not imply causality: that is, that the presence of the technology would reduce the likelihood of having negative evaluations of operational characteristics or the overall assessments. Instead, they present the difference in means conditional on mention of a technology between the two subgroups. Many other factors, as we describe later in this paper, are expected to contribute to the difference.

3. Empirical approach

3.1. Specifications and estimation

The content analysis data were used to build a linear probability model (LPM) exploring the relationship of fuel-saving technologies with the various operational characteristics. Following Helfand et al. (2016), we run the following LPM as our baseline model predicting $I(NegativeOper)_{i,j,t}$, an indicator variable equal to 1 if operational characteristic j was negatively reviewed on model-year t vehicle in auto review i:

$$I(NegativeOper)_{i,j,t} = \sum\nolimits_k \beta_k I(Tech)_{i,k,t} + FixedEffects + \epsilon_{i,j,t} \tag{1}$$

in which $I(Tech)_{i,k,t}$ is a vector of k indicator variables representing all fuel-saving technologies examined in this study. The indicator variable is equal to 1 if a technology was mentioned in an auto review. *FixedEffects* include, at a minimum, website, class, and make fixed effects, to address potential unobserved heterogeneity in factors that might be correlated with both the technology mentioned and the operational characteristic.

When we pool the MY2014 and 2015 data, we also control for the following fixed effects: year (e.g., market conditions common to all manufacturers), year-by-website (e.g., a website's year-specific review standards and preferences), year-by-class (e.g., year-specific market conditions for a vehicle class common to all manufacturers), and year-by-make (e.g., a company's year-specific innovation and/or production strategy). The interactions of fixed effects play important roles in identifying our variables of interest β_k, as they control for factors that do not vary over the make-year, class-year, and website-year. Model-by-class fixed effects interacting with model year will also be included in our robustness check, although within model-by-class-by-year variation in the variables of fuel-saving technologies could be reduced.

We estimate this specification using a standard fixed effects regression. A positive coefficient of β_k indicates that the mentioned technology is associated with an increased likelihood of a hidden cost; a negative coefficient, on the other hand, indicates that the technology is associated with a reduction in the likelihood of a hidden cost.

One hypothesis suggested in Helfand et al. (2016) is that, while it appears possible for all the technologies to be used without imposing hidden costs, problems may arise due to variation in quality of implementation of some technologies in some vehicle models. In that paper, the authors compare effects on operational characteristics when the technology is mentioned, to effects on characteristics when the technology receives a negative evaluation; they find more correlations with negatively reviewed

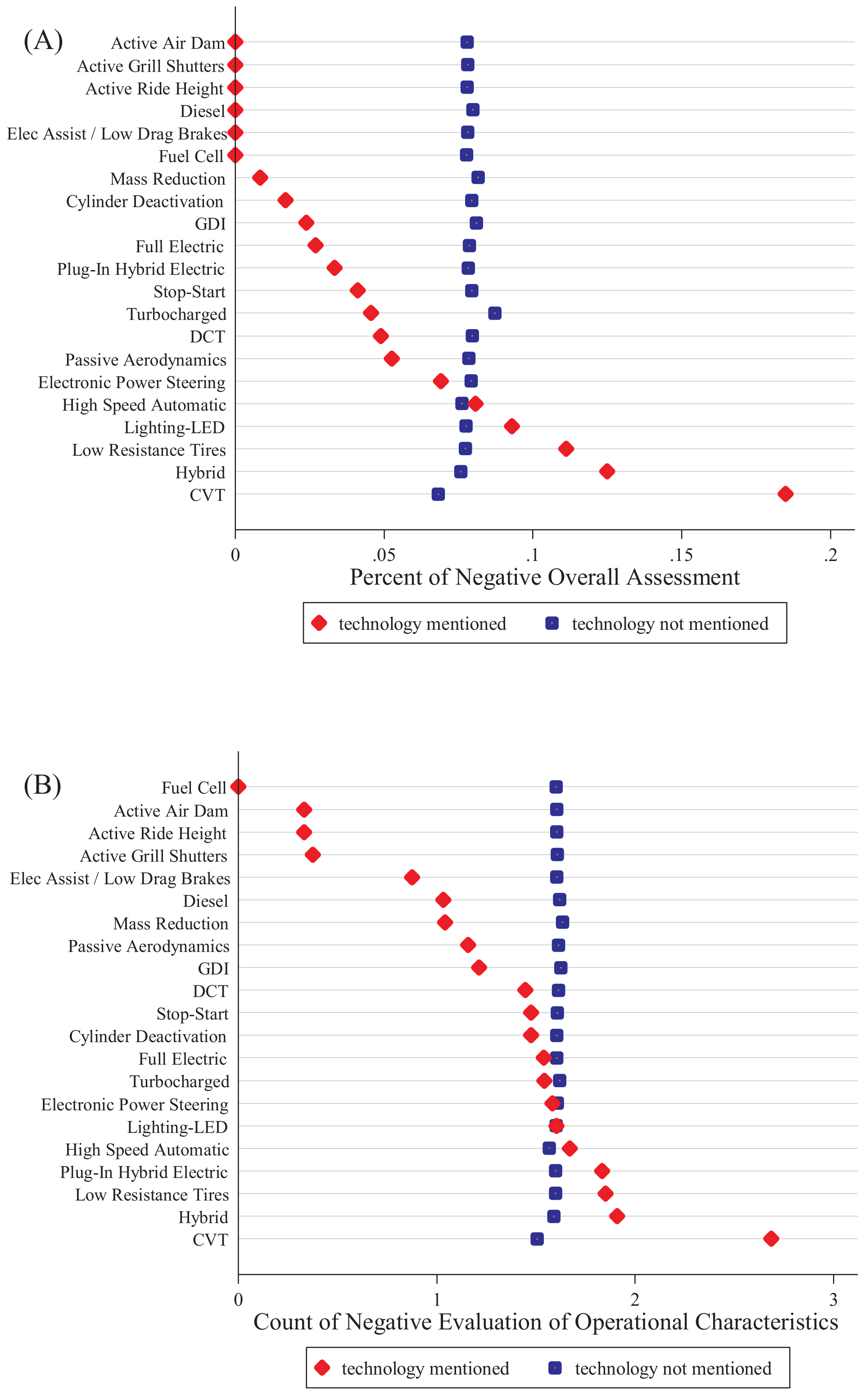


Fig. 2. Comparison of evaluation of overall assessment and operational characteristics by technology using pooled data.

characteristics for negatively reviewed technologies than for the presence of the technologies. In this paper, we directly estimate whether negative operational impacts are responsive to a negatively reviewed technology conditional on the presence of the technology[2]:

$$I(NegativeOper)_{i,j,t} = \sum_k \alpha_k I(NegativeTech)_{i,k,t} + \sum_k \beta_k I(Tech)_{i,k,t} + FixedEffects + \epsilon_{i,j,t} \quad (2)$$

$I(NegativeTech)$ equals 1 if fuel-saving technology k was negatively reviewed in auto review i of model year t, and equals 0 if technology k was positively or neutrally reviewed, or not mentioned. The underlying idea is that, if a technology was not implemented well (e.g., poor quality in production of the technology, poor installation, and/or poor other adjustments to the technology), auto reviewers would give a negative review for the technology.[3]

Conditional on the mention of fuel-saving technologies and the fixed effects, this specification directly tests whether a negatively reviewed technology is correlated with negative rating of an operational characteristic. Also, specification (2) seeks to address the selection bias arising from the possibility that a negatively reviewed technology is more likely to be mentioned than a technology that is working well. The coefficients of interest, β_k and α_k, are estimates of the increase in the probability of a negative review due to either the presence of technology k or a negative evaluation of technology k, respectively. A positive value for α_k, especially if combined with a reduction in the value for β_k, supports the hypothesis that poor implementation of technologies is associated with negative ratings of operational characteristics, perhaps more than the mere presence of the technologies.

Lastly, we replace operational characteristics on the left hand side of specification (2) by using the overall vehicle assessment to estimate its relationship with fuel-saving technologies.

3.2. Potential estimation concerns

Including the fixed effects is useful for identifying the relationship between fuel-saving technologies and reviewed operational characteristics by addressing a variety of potential confounders. Yet, we recognize that there remains a possible selection bias. For instance, if some technologies were put into low-quality vehicles, and the low quality is not fully controlled for by the fixed effects, the estimated coefficient would be biased upward. As a result, we describe the results as correlations between technologies and operational characteristics rather than claiming causality. We thus focus on whether statistically significant positive coefficients are consistently estimated from our regression models.

It is important to note that absence of a mention in a review does not mean that the technology is absent; it means that the reviewer did not mention it. It is plausible that auto reviewers would notice and comment on undesirable features more than on positive features. If so, the estimated relationship between a technology and a hidden cost may be biased upward.

In addition, one concern with LPM is that estimated probability may be not bounded between -1 and 1. However, with binary dependent variables, LPM has some advantages over logit models in that causal analysis is valid and does not require functional form assumptions about the error term (Angrist and Pischke, 2009). In fact, almost all of our estimated coefficients in this study are between -1 and 1.

Last, while robust standard errors are generally used for non-constant error variance with LPM, they are subject to small sample bias and high sampling variance (Angrist and Pischke, 2009, p. 307). As a result, robust standard errors may be too small by accident and thus increase rejection of the null hypothesis. In this study, in additional to using conventional robust standard errors, we follow the suggestion of Angrist and Pischke (2009) of using the maximum of the conventional standard error and a robust standard error as our best measure of precision.

4. Estimated relationship of technologies and operational characteristics

This paper focuses on examining whether the estimated relationships using multiple datasets (MY 2014, MY 2015, and pooled) with a rich set of fixed effects are consistent and robust. For each dataset, we run a separate regression of each of 22 operational characteristics on all fuel-saving technologies examined and a set of fixed effects. There are 20 technologies for MY 2014 data and 21 for MY 2015 data. Thus, we have 440 estimated coefficients for the fuel-saving technologies for MY 2014 data (20 coefficients for each operational characteristic), and 462 estimated coefficients for MY 2015 and for the pooled data (21 coefficients for each operational characteristic).

In this section, first, we present the estimated results of the initial specification, separately for each model year. Second, we report detailed results using the pooled data, with its advantage of more observations as well as additional fixed effects (i.e., website-year, class-year, and make-year). This section also includes robustness checks.

4.1. Overview of the results across datasets

Fig. 3 provides the number of significant coefficients in the 22 regressions across alternative datasets based on significance at the 10% level, and use of the measure for standard errors suggested by Angrist and Pischke (2009); the detailed estimated results are

[2] To be precise, the presence of the technology is assumed from its being mentioned in the review, rather than from independent identification of its presence in the specific vehicle.

[3] In this study, a technology would be coded as negative when a passage of text contained a negative evaluation of the technology in an auto review. For instance, CVTs would be coded as negative for "the CVT isn't particularly responsive (Autotrader 2015)," "the CVT keeps the engine droning away at high revs to make any sort of power (Motortrend, 2014)," or "should you opt for the CVT, the trio of cylinders will grumble in protest every time you try to accelerate (Caranddriver, 2014)."

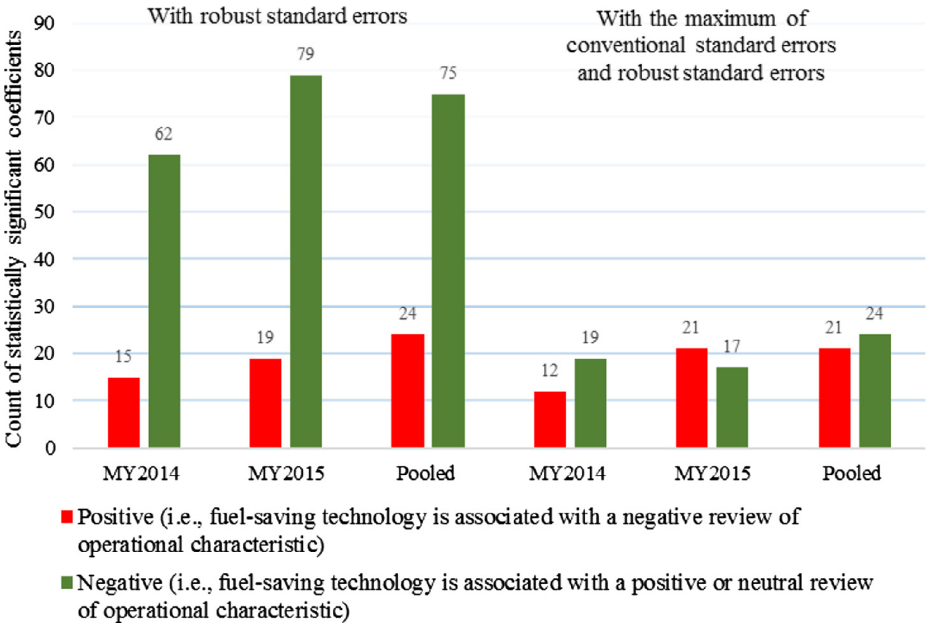


Fig. 3. Overview of estimation results using alternative datasets and standard errors (10% significance level).

reported in Appendix Table A.1 and Table A.2 for MY 2014 and MY 2015 vehicles, respectively. Recall that a positive coefficient indicates that fuel-saving technology is associated with a negative review of operational characteristic. As Fig. 3 shows, only 2.7% (12 out of 440) of coefficients for the

MY 2014 data, and 4.5% (21 out of 462) of coefficients for both the MY 2015 and the pooled data are positive and statistically significant. The results for positive relationships, our focus, do not seem to be sensitive to the standard errors we use. The general pattern of MY 2014 results, of few positive and statistically significant coefficients, continues for MY 2015.

Only four of the 12 coefficients, or about 1% of the 440 coefficients for MY 2014, that are positive and statistically significant using MY 2014 data remain positive and statistically significant using MY 2015 data: hybrid is associated with a negative rating for brake feel, plug-in hybrid electric is associated with a negative rating for powertrain noise, and CVT is associated with negative ratings for general drivability and powertrain noise. The small number of consistently significant associations between technologies and negatively reviewed characteristics raises the question whether they are significant by chance. On the other hand, in a single model-year, the small sample sizes for some technologies may make statistically significant relationships difficult to detect.

Except with the maximum of the conventional standard error and a robust standard error for the MY 2015 data, there appear to be more statistically significant negative coefficients than positive ones: that is, there may be more cases of hidden benefits than hidden costs. It is noteworthy that the number of negative relationships is substantially reduced (from over 60 to under 20 in either MY 2014 and MY 2015 reviews) using the approach suggested by Angrist and Pischke compared to the estimated results using robust standard errors. This observation demonstrates how this approach can affect interpretation of results by creating a more stringent standard for significance.

4.2. Results of pooled data

The results from estimating specification (1) with the pooled data from MY 2014 and MY 2015 include year, year-by-website, year-by-class, and year-by-make fixed effects, in addition to the fixed effects from the individual-year analyses. The estimation results for all operational characteristics are summarized in Fig. 3 and detailed in Table 6.

Similar to the results with single-year datasets, there continue to be relatively few cases of fuel-saving technologies correlated with a negative rating for an operational characteristic, i.e., positive coefficients. Out of 462 coefficients, 21 coefficients are positive and statistically significant using the measure of precision suggested by Angrist and Pischke; with robust standard errors as our measure of precision, 24 coefficients are statistically significant and positive. In addition, the magnitudes are small; only 5 of the 462 coefficients are associated with an increased probability of a negative impact of more than 0.15.

Among the 21 positive and significant coefficients, five coefficients associate CVT with negative evaluations, for cornering ability, general drivability, acceleration capability, wind noise, and powertrain noise. These coefficients are fairly small: the maximum estimated increased probabilities are 0.14, for acceleration capability and powertrain noise. Plug-in hybrid vehicles have three of the largest positive significant coefficients, for brake feel (0.21), range (0.22), and powertrain noise (0.4).

We note again that the correlations may be affected by unobserved variables. For instance, it is possible that, instead of CVT itself, lower quality implementation of CVT contributes to the negative associations with negative rating of operational characteristics. Also, it is possible that CVTs were put into vehicles with a relatively loud powertrain noise or other problems that are not captured by

Table 6
Relationships between the presence of fuel-saving technologies and negatively reviewed operational characteristics, pooled MY 2014 and 2015 data.

	(1) Negative Steering Feel	(2) Negative Cornering Ability	(3) Negative General Drivability	(4) Negative General Handling	(5) Negative Acceleration Feel	(6) Negative Acceleration Capability	(7) Negative General Acceleration	(8) Negative Brake Feel	(9) Negative Stopping Ability	(10) Negative General Braking	(11) Negative Tire-Road Noise
Active Air Dam	−0.00	−0.11	−0.02	0.23	0.01	−0.13	−0.06	0.03	−0.01	0.01	0.04
Active Grille Shutters	−0.10*	−0.12***	−0.09**	−0.05	0.10	−0.12	−0.02	−0.04	−0.05**	−0.02*	−0.08*
Active Ride Height	−0.24**	−0.04	−0.01	−0.03	−0.16***	0.17	0.00	−0.02	0.01	0.02	0.01
Low Resistance Tires	0.01	0.14	0.02	0.01	−0.02	−0.08	−0.00	**−0.09****	0.05	0.02	0.06
Electronic Power Steering	**0.11*****	0.03	0.01	−0.02	−0.00	0.01	−0.01	0.01	**0.02***	−0.01	**0.04***
Turbocharged	0.03	0.00	0.00	−0.01	−0.00	0.04*	−0.01	0.01	0.00	−0.01	0.03
GDI	−0.03	**−0.06*****	−0.02	−0.02	−0.00	0.03	−0.01	−0.03	0.01	−0.00	−0.00
Cylinder Deactivation	0.07	0.02	−0.02	−0.05	0.03	0.01	0.04	**0.09***	0.02	−0.02*	−0.02
Diesel	0.03	−0.01	−0.04	−0.01	−0.01	**0.12***	**0.06***	−0.02	0.01	−0.02*	−0.02
Hybrid	0.02	−0.06*	0.05	−0.02	0.05	0.01	0.02	**0.12*****	−0.01	0.01	−0.05*
Plug-In Hybrid Electric	−0.00	0.05	−0.01	0.05	−0.03	−0.01	0.01	**0.21****	0.00	0.02	0.02
Full Electric	0.04	−0.04	0.08	−0.04	0.04	0.10	0.03	0.08	0.01	0.01	−0.05
Stop-Start	−0.01	**−0.06****	−0.06*	**−0.06***	−0.03	**−0.09****	−0.00	−0.05	−0.00	−0.01	−0.02
High Speed Automatic	−0.02	0.00	**−0.03****	−0.00	−0.01	**−0.04****	−0.01	0.01	0.00	−0.01	0.00
CVT	0.02	**0.07****	**0.11****	0.04	0.02	**0.14*****	−0.00	0.01	0.03	0.02	0.02
DCT	0.01	**0.07****	−0.00	0.02	0.01	−0.02	0.00	0.02	0.00	−0.01	**0.06***
Elec Assist/Low Drag Brakes	−0.03	−0.08*	−0.07*	−0.05	0.02	−0.04	−0.05**	−0.07**	0.05	−0.01	−0.04
Lighting-LED	−0.06	−0.04	0.04	0.06	−0.07*	−0.06	0.00	−0.04	−0.03	−0.02	0.09
Mass Reduction	0.02	−0.00	**−0.06*****	−0.01	**−0.05*****	−0.03	−0.02	0.04	−0.02	0.00	−0.01
Passive Aerodynamics	−0.05	−0.03	−0.05	−0.01	−0.00	−0.01	0.02	0.05	−0.02	0.03	0.01
Fuel Cell	−0.18***	−0.17***	−0.29***	−0.04	−0.11*	−0.32***	−0.03	−0.02	−0.05	0.02	−0.06

	(12) Negative Wind Noise	(13) Negative Interior Noise	(14) Negative Powertrain Noise	(15) Negative General Noise	(16) Negative Chassis Vibration	(17) Negative Powertrain Vibration	(18) Negative General Vibration	(19) Negative Ride Comfort	(20) Negative Fuel Economy	(21) Negative Range	(22) Negative Charging
Active Air Dam	−0.08	0.01	−0.06	−0.09	−0.01	−0.01	0.01	0.02	0.03	0.00	0.00
Active Grille Shutters	−0.01	0.00	−0.13**	−0.02	0.00	−0.01	−0.00	−0.11**	−0.03	−0.01	−0.00
Active Ride Height	0.02	−0.01	−0.23***	0.02	0.00	−0.01	−0.00	0.01	−0.20**	0.01	0.01
Low Resistance Tires	−0.04**	−0.01	−0.00	0.05	−0.00	0.03	0.03	−0.03	−0.11**	0.04	−0.01
Electronic Power Steering	0.00	−0.00	0.00	**−0.02***	0.01	0.00	−0.00	−0.00	0.01	−0.00	−0.00
Turbocharged	0.01	**0.01***	0.03	−0.02	0.00	**0.01****	0.00	−0.01	0.02	−0.01**	−0.00
GDI	0.01	−0.01	0.01	0.03	−0.01	−0.00	0.01	0.01	−0.05	0.00	0.00
Cylinder Deactivation	−0.04***	−0.01	0.01	−0.03	**−0.02***	0.00	0.01	0.09	**−0.15*****	0.00	−0.00
Diesel	−0.01	0.01	−0.05	0.01	−0.00	0.01	−0.01	**−0.11****	**−0.09***	−0.00	0.00
Hybrid	−0.00	−0.01	0.02	−0.00	−0.00	−0.01*	−0.01*	−0.01	−0.00	−0.01**	−0.00

(continued on next page)

Table 6 (*continued*)

	(1) Negative Steering Feel	(2) Negative Cornering Ability	(3) Negative General Drivability	(4) Negative General Handling	(5) Negative Acceleration Feel	(6) Negative Acceleration Capability	(7) Negative General Acceleration	(8) Negative Brake Feel	(9) Negative Stopping Ability	(10) Negative General Braking	(11) Negative Tire-Road Noise
Plug-In Hybrid Electric	−0.01	0.01	**0.40*****	−0.02	−0.00	0.00	0.05	−0.04	0.07	**0.22*****	0.05
Full Electric	0.00	−0.02*	**−0.13****	−0.05	−0.00	−0.02	−0.02*	−0.04	**−0.13*****	0.02	0.02
Stop-Start	**−0.06*****	−0.02*	−0.02	0.00	−0.01	0.01	0.02	−0.04	−0.05	0.01	−0.00
High Speed Automatic	**−0.02***	−0.01	0.01	**−0.04*****	0.00	−0.00	−0.00	0.01	−0.02	**−0.01*****	−0.00*
CVT	**0.05****	0.01	**0.14*****	0.02	0.01	0.00	0.01	**−0.09****	−0.07	−0.01	−0.00
DCT	−0.02	0.02	0.01	−0.00	0.01	0.00	0.01	**0.09****	−0.01	−0.01	−0.00
Elec Assist/Low Drag Brakes	−0.03	−0.01	0.03	−0.07*	−0.01	−0.01	−0.01	−0.01	−0.11*	0.12	−0.01
Lighting-LED	**−0.05*****	0.05	−0.05	−0.04	0.01**	−0.01	0.00	0.00	−0.06	−0.00	0.00
Mass Reduction	−0.01	−0.01	−0.01	−0.02	−0.01	0.00	0.00	**0.10*****	−0.06**	0.01	0.01
Passive Aerodynamics	0.02	0.01	**−0.08****	−0.03	−0.01	−0.02**	**−0.03*****	−0.06	−0.01	−0.00	−0.00
Fuel Cell	0.00	0.01	0.02	0.05*	−0.00	−0.00	−0.04	−0.19***	−0.42***	−0.01	−0.00

Notes: Number in a cell indicates the estimated probability of negative rating for the column variable conditional on technology named in the row. Number in bold means the significant result holds when we use the maximum of robust standard errors and conventional standard errors as the best measure of efficiency, instead of using robust standard errors. For all columns, the sample size is 2,238 from the pooled data of model years 2014 and 2015. For each column, the linear probability model regresses the column variable on all 21 technologies and a set of fixed effects, including year, website, vehicle class, vehicle make, website-by-year, class-by-year, and make-by-year. Asterisks indicate the level of statistical significance: 10% (*), 5% (**), and 1% (***) levels.

Table 7
Estimated Relationships (α_k) between negatively reviewed fuel-saving technologies and negatively reviewed operational characteristics, pooled MY 2014 and 2015 data.

	(1) Negative Steering Feel	(2) Negative Cornering Ability	(3) Negative General Drivability	(4) Negative General Handling	(5) Negative Acceleration Feel	(6) Negative Acceleration Capability	(7) Negative General Acceleration	(8) Negative Brake Feel	(9) Negative Stopping Ability	(10) Negative General Braking	(11) Negative Tire-Road Noise
Active Air Dam	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Active Grille Shutters	−0.06	0.05	0.14	−0.04	−0.12	−0.01	−0.05	0.16***	−0.02	0.05**	0.11*
Active Ride Height	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Low Resistance Tires	0.12	0.33	**0.41****	0.21	0.17	0.05	0.14	−0.05	−0.04	−0.08	−0.06
Electronic Power Steering	**0.80*****	**0.14****	**0.10****	0.04	0.02	0.03	−0.03*	0.03	0.05	0.01	0.04
Turbocharged	**0.12****	**0.13****	**0.15****	−0.01	**0.27*****	**0.34*****	0.03	0.01	0.03	−0.00	0.03
GDI	0.14	−0.08**	−0.14**	−0.16***	0.16	0.14	−0.03	−0.02	−0.08**	−0.01	0.10
Cylinder Deactivation	0.18	0.11	0.15	**0.38***	−0.02	−0.15*	−0.09**	**−0.19*****	−0.06	−0.00	0.18
Diesel	−0.09	−0.01	0.04	0.08	0.09	**0.45*****	**0.27****	−0.03	0.07	−0.01	−0.07
Hybrid	0.02	−0.00	0.17	0.02	0.07	**0.22****	0.05	0.11	−0.01	0.05	−0.02
Plug-In Hybrid Electric	0.06	0.03	−0.13	0.01	0.05	−0.04	0.03	−0.22	0.15	−0.03	−0.00
Full Electric	0.47	−0.06	−0.10	0.02	0.51	0.37	**0.52***	0.48	−0.12	−0.05	0.01
Stop-Start	0.05	−0.11**	0.10*	0.06	0.12	−0.04	−0.01	−0.02	0.08	0.03	**0.14***
High Speed Automatic	**0.10*****	**0.08****	**0.07****	**0.09*****	**0.19*****	**0.15*****	**0.03***	**0.06****	0.02	0.00	0.03
CVT	0.05	0.05	**0.17****	**0.17*****	**0.14****	**0.33*****	0.03	−0.00	0.00	**−0.04***	0.05
DCT	−0.03	0.12	**0.24*****	0.07	**0.15****	0.02	0.04	0.08	0.08	0.00	0.09
Elec Assist/Low Drag Brakes	−0.16	0.01	−0.15*	−0.02	**0.99*****	**0.90*****	0.02	0.02	−0.05	−0.01	0.03
Lighting-LED	−0.13	−0.17**	−0.25***	−0.28***	−0.06	−0.10	0.01	−0.08	−0.13	0.01	−0.27***
Mass Reduction	**0.76*****	0.19	−0.13	**0.65****	0.03	**−0.37*****	0.01	−0.11**	−0.07	0.01	−0.04
Passive Aerodynamics	0.01	−0.06	−0.08	0.31	**0.39****	**0.31***	**0.34***	**−0.18*****	−0.04	−0.05	0.08
Fuel Cell	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	(12) Negative Wind Noise	(13) Negative Interior Noise	(14) Negative Powertrain Noise	(15) Negative General Noise	(16) Negative Chassis Vibration	(17) Negative Powertrain Vibration	(18) Negative General Vibration	(19) Negative Ride Comfort	(20) Negative Fuel Economy	(21) Negative Range	(22) Negative Charging
Active Air Dam	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Active Grille Shutters	0.01	−0.00	−0.02	0.03	−0.01	0.02	−0.04	−0.06	−0.11	0.00	0.00
Active Ride Height	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Low Resistance Tires	0.04	−0.02	0.14	0.18	−0.02	−0.08	0.09	**0.35***	−0.03	−0.01	−0.01
Electronic Power Steering	0.01	−0.01	−0.03	**0.07***	0.02	−0.01	−0.01	−0.00	0.02	0.01	0.00
Turbocharged	−0.02	−0.02	**0.15*****	0.01	0.02	0.03	−0.00	−0.01	**0.16****	0.00	0.00*
GDI	−0.00	−0.04**	0.19	0.00	0.00	−0.01	**0.21***	−0.04	0.02	−0.01	−0.00
Cylinder Deactivation	0.02	0.01	0.30	0.04	0.02	0.00	−0.05	−0.03	0.26	0.01	−0.00
Diesel	−0.06	−0.03	**0.29****	−0.04	−0.00	−0.03*	0.02	−0.03	0.18	−0.00	−0.01
Hybrid	0.01	−0.04	0.13	0.03	0.00	0.00	−0.00	0.06	0.15	0.01	−0.00
Plug-In Hybrid Electric	−0.02	0.01	0.25	−0.06	0.01	0.01	0.14	0.01	0.16	**−0.27*****	0.08

(continued on next page)

Table 7 (*continued*)

	(1) Negative Steering Feel	(2) Negative Cornering Ability	(3) Negative General Drivability	(4) Negative General Handling	(5) Negative Acceleration Feel	(6) Negative Acceleration Capability	(7) Negative General Acceleration	(8) Negative Brake Feel	(9) Negative Stopping Ability	(10) Negative General Braking	(11) Negative Tire-Road Noise
Full Electric	0.05	0.04	0.26	**0.65****	−0.02	−0.01	−0.06	0.01	−0.05	**1.00*****	−0.06
Stop-Start	0.06	−0.01	−0.02	0.05	−0.00	−0.01	0.04	**0.20****	0.05	−0.02	0.00
High Speed Automatic	0.02	**0.03***	**0.07****	−0.00	−0.01*	**0.02***	−0.00	**0.13*****	**0.12*****	0.00	−0.00
CVT	**0.10****	0.01	**0.26*****	0.06	0.01	0.01	0.03	**0.15****	**0.14****	0.00	−0.00
DCT	−0.00	0.08	−0.02	0.04	−0.01	0.03	0.07	**0.20****	−0.01	0.00	−0.00
Elec Assist/Low Drag Brakes	−0.06	−0.01	−0.07	0.04	0.05*	0.02	0.02	0.03	0.21**	−0.12	−0.07
Lighting-LED	−0.06	−0.07	−0.06	**0.81*****	0.03**	0.01	0.01	−0.46***	0.07	0.02	0.00
Mass Reduction	−0.00	−0.00	0.50	0.31	−0.01	−0.02	−0.01	0.45	0.11	−0.01	−0.00
Passive Aerodynamics	0.24	0.13	0.01	0.01	0.01	−0.02	0.01	−0.10	**0.38***	−0.00	−0.01
Fuel Cell	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes: Number in a cell indicates the estimated probability of negative rating for the column variable conditional on technology named in the row. Number in bold means the significant result holds when we use the maximum of robust standard errors and conventional standard errors as the measure of efficiency, instead of using robust standard errors. For all columns, the sample size is 2,238. For each column, the linear probability model regresses the column variable on all technologies that were negatively reviewed and a set of fixed effects, including year, website, vehicle class, vehicle make, website-by-year, class-by-year, and make-by-year. Asterisks indicate the level of statistical significance: 10% (*), 5% (**), and 1% (***) levels.

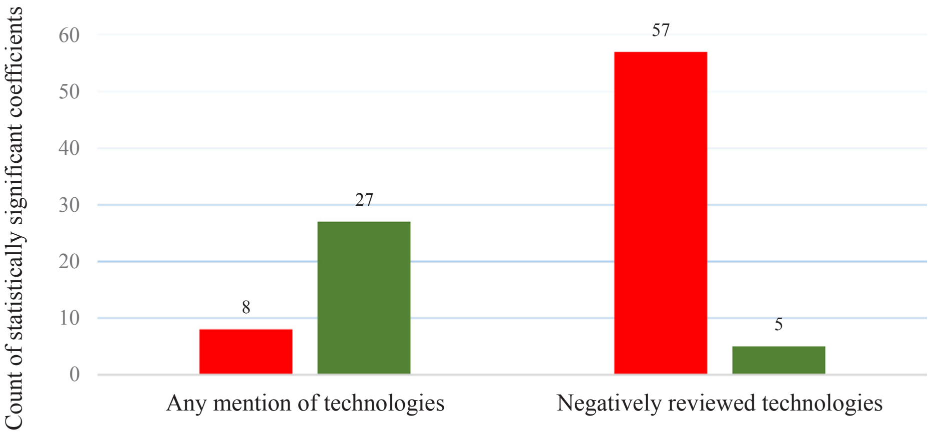


Fig. 4. Overview of estimation results for variation in implementation quality (10% significance level).

the fixed effects. We consider these concerns in the next two subsections.

4.2.1. *Variation in implementation quality for fuel-saving technologies?*

The analyses above are based on the presence of the technology; negative effects associated with the presence of the technology may be due to an inherent property of the technology. In contrast, specification (2) seeks to distinguish between problems that are inherent to a technology, and problems associated with particular use, installation of the technology, or other adjustments made to the vehicle with the technology. We do so by including variables for both the presence of the technology (β_k) and for a negatively reviewed technology (α_k).

Complete estimated results of α_k, the coefficient for negatively reviewed technology k, are reported in Table 7, while complete estimated results of β_k, the coefficient for any mention of technology k, are in Appendix Table A.3. Summarized in Fig. 4, the results using the pooled data and the Angrist and Pischke measure of precision find 57 out of 462 $\alpha_k s$ are positive and statistically significant, while five coefficients are negative and statistically significant. In addition, for β_k, eight out 462 are positive and statistically significant, substantially less than the 21 positive coefficients from estimating specification (1). Because a positive coefficient indicates a fuel-saving technology (either mentioned (β_k) or negatively reviewed (α_k)) is associated with a negative review of an operational characteristic, the results suggest that negatively reviewed technologies, instead of the presence of the technologies themselves, are more likely to be associated with negative ratings of operational characteristics. The results suggest that vehicles that did not get negative evaluations on fuel-saving technologies may have been able to implement the technologies without harm to operational characteristics. We repeat, though, that these data are not sufficient to demonstrate causality. For instance, it is possible that technologies are negatively reviewed most often in the context of a negatively reviewed operational characteristic.[4]

Using CVT as an example, Table 7 shows that eight of the 57 positive coefficients (and one of the 21 negative coefficients) of α_k are with CVT, and there are no positive coefficients (but two negative coefficients) of β_k with CVT. One explanation for this finding is that poor implementation of CVT (from the negative reviews of CVT in certain models, α_k), rather than CVT itself (from mention of the technology, β_k), is related to negative rating of operational characteristics. Recall that specification (1) found that the presence of CVT was associated with negative ratings for five operational characteristics; specification (2) suggests that negatively evaluated CVTs may contribute to negative ratings related to drivability, acceleration, noise, ride comfort, and fuel economy, as shown in Table 7.

Similarly, negatively reviewed high speed automatic and turbocharging technologies are associated with negative ratings of 13 and 7 operational characteristics, respectively, while the presence of the two technologies shows little relationship to negatively reviewed operational characteristics based on both specifications (1) and (2). This observation suggests that specification (2) may provide a clearer signal of the difference between the effects of the presence of a technology, and the effects of a poorly reviewed technology, on operational characteristics.

These results are consistent with Helfand et al. (2016)'s proposal that quality of implementation, rather than technologies

[4] For instance, one coded segment reads, "With the CVT and direct-injection engine technology new to the 2015 CR-V, some owners have reported experiencing a vibration through the driver's seat, and unfortunately I'm among them. The subtle vibration is intermittent, but when it happens, it does so while the crossover is idling" (Gale, 2015). This is coded negative for CVT, for GDI, and for general vibration. Here, the CVT is rated negatively due to its association with vibration.

themselves, is associated with vehicle qualities.

4.2.2. Potential model-by-class-by-year specific unobserved vehicle attributes

Another potential concern that could affect our findings is that fuel-saving technology put into vehicles might be systematically related to other attributes or issues that are unobserved by researchers. To assess the concern, we adjust specification (2) by using model-by-class fixed effects interacted with model year, instead of using make-by-year and class-by-year fixed effects.[5] Coefficients of α_k and β_k are identified within model-class-year variation in whether the fuel-saving technologies we examine are present or not. These fixed effects control for all vehicle characteristics that stay constant for the model-class in a model year. For example, vehicle characteristics in the error term in the previous specifications that are common to the Toyota Highlander standard SUVs in MY 2014 and would affect the operational characteristics will be controlled by these fixed effects separately from vehicle characteristics common to Toyota Highlander small SUVs.

The inclusion of these fixed effects still does not completely address the potential for selection bias. For instance, if a CVT were put into a trim that is already noisy, our estimates of β_k would be biased upward for the relationship between powertrain noise and the technology. In addition, these fixed effects will reduce the variation of fuel-saving technologies for some models. For instance, there is no variation of hybrid technology within the model Toyota Prius C; the estimated results would be based on other models adopting hybrid technology. We present the analysis for the purpose of a robustness check. If the results show a different pattern of the relationships between fuel-saving technologies and operational characteristics, our findings above may be affected by selection bias.

Using model-by-class fixed effects interacting with model year, estimated results of adjusted specification (2) are similar to the results in subsection 0. Here, 41 out of 462 coefficients of the negatively reviewed technologies (α_k) are positive and statistically significant, as reported in Table A.4, compared to 57 with the more limited fixed effects (in Table 7); as in the previous specification, 5 out of 462 coefficients of the presence of the technologies (β_k) are positive and statistically significant, as reported in Table A.5. Negatively rated CVT, turbocharging, and high speed automatic show a pattern of relationships with negative ratings of operational characteristics consistent with the results in Table 7.

In sum, controlling for the model-class-year-specific effects provides a similar pattern of relationships between fuel-saving technologies and operational characteristics. That is, positive and significant relationships between the presence of fuel-saving technologies and negative rating of operational characteristics are a small proportion of the possible relationships, and are outnumbered by the 18 negative and significant relationships. Also, we continue to find that negatively reviewed technologies, instead of the presence of the technologies themselves, are more likely to be associated with negative ratings of operational characteristics. The results suggest that there might be no serious selection bias, though we still cannot rule out a potential bias arising from the possibility that the presence of a new technology is endogenous for a trim within a vehicle model.

5. Relationship of technologies and overall assessment of the vehicle

We have shown in the previous section that the presence of fuel-saving technologies is infrequently associated with negative evaluations of operational characteristics; when there is a relationship, it is slightly more likely negative, implying hidden benefits, than positive. Although some technologies show a positive relationship with several negative operational characteristics, problems with implementation of the technologies may be a better explanation for the relationship than the existence of the technology.

It may also be useful to know how the technologies are associated with the overall summary evaluation of a vehicle – the recommendation whether a vehicle is worth buying. Auto reviewers' overall rating is usually highlighted after consideration of vehicle characteristics and comparison with similar vehicles in the same vehicle segment. It is expected to summarize all positive and negative impacts on vehicle quality, including any impacts not separately addressed in the review. The overall rating is expected to matter for vehicle buyers' decisions, as it provides a recommendation about whether it is reasonable or not to purchase the vehicle among the models sharing similar features in the market.

In this section, we investigate the relationship between the overall rating and fuel-saving technologies, using a coded variable denoting that the overall assessment provided by a reviewer is positive, negative, or mixed. We begin by substituting the left hand side of specification (2) with $I(NegativeOverall)$, an indicator variable for whether the vehicle reviewed got a negative overall assessment, to obtain the following specification:

$$I(NegativeOverall)_{i,t} = \sum\nolimits_k \alpha_k I(NegativeTech)_{i,k,t} + \sum\nolimits_k \beta_k I(Tech)_{i,k,t} + FixedEffects + \epsilon_{i,t} \tag{3}$$

Columns (1a) and (1b) of Table 8 report the results of estimating specification (3) using the maximum of conventional and robust standard errors as our measure of precision. We do not find evidence that the presence of fuel-saving technologies is positively related to negative overall assessments, as shown in (1a). Instead, we find that the presence of three technologies – turbocharged, cylinder deactivation, and high speed automatic – is associated with a reduced likelihood of a negative overall assessment. In addition, all estimated coefficients are less than 0.1 in absolute value no matter whether they are statistically significant. The fairly small coefficients indicate that the overall rating of a vehicle appears not to be responsive to the presence of the fuel-saving technologies.

Column (1b) of Table 8 shows six negatively reviewed technologies – low resistance tires, CVT, electronic power steering, hybrid, high speed automatic, and turbocharged – are significantly correlated with negative overall assessments. The results again raise the possibility that poorly implemented technologies (1b) in some vehicle models, instead of the technologies themselves (1a), are

[5] We use only specification (2) because specification (1) is nested in it.

Table 8
Relationship between the presence of a fuel-saving technology, a negatively reviewed technology and operational characteristic, and the overall assessment of the vehicle.

Technology	(1a) Presence of Technology	(1b) Negative Review of Technology	(2a) Presence of Technology	Operational Characteristics	(2b) Negative Review of Operational Characteristics
Active Air Dam	−0.02	0.00	0.00	Steering Feel	0.06***
Active Grill Shutters	−0.08	−0.02	0.00	Cornering Ability	−0.01
Active Ride Height	−0.02	0.00	−0.02	General Drivability	0.15***
Low Resistance Tires	−0.03	0.29*	0.08	General Handling	0.09***
Electronic Power Steering	−0.02	0.08**	−0.01	Acceleration Feel	0.07***
Turbocharged	−0.04***	0.12**	−0.03*	Acceleration Capability	0.08***
GDI	−0.03	0.02	−0.03	General Acceleration	0.06
Cylinder Deactivation	−0.07*	−0.06	−0.07*	Brake Feel	0.09***
Diesel	−0.05	−0.08	−0.06*	Stopping Ability	0.02
Hybrid	−0.04	0.21**	0.00	General Braking	−0.02
Plug-In Hybrid Electric	0.03	−0.05	−0.04	Tire−Road Noise	0.03
Full Electric	−0.04	0.03	−0.05	Wind Noise	0.02
Stop-Start	−0.03	−0.00	−0.01	Interior Noise	0.07
High Speed Automatic	−0.04***	0.19***	−0.00	Powertrain Noise	0.06***
CVT	−0.05	0.23***	0.01	General Noise	0.02
DCT	−0.00	0.07	−0.00	Chassis Vibration	−0.02
Elec Assist Or Low Drag Brakes	−0.06	0.08	−0.05	Powertrain Vibration	0.11*
Lighting-LED	−0.00	−0.09	−0.01	General Vibration	−0.05
Mass Reduction	−0.04	−0.16	−0.04	Ride Comfort	0.04**
Passive Aerodynamics	−0.02	0.16	0.00	Fuel Economy	0.07***
Fuel Cell	−0.06	0.00	0.05	Range	−0.01
				Charging	0.00

Notes: There are 2238 observations for the two specifications. Dependent variable for the two specifications is an indicator variable for negative overall assessment. Independent variables of the first specification include the presence of the technologies (column 1a) plus negatively reviewed technologies (column 1b). Independent variables of the second specification include the presence of the technologies (column 2a) plus negatively reviewed operational characteristics (column 2b). Number in a cell indicates the estimated coefficient. All specifications include a set of fixed effects, including year, website, vehicle class, vehicle make, year-by-website, year-by-class, and year-by-make. Asterisks indicate the level of statistical significance using the maximum of robust standard errors and conventional standard errors as the measure of efficiency: 10% (*), 5% (**), and 1% (***) levels.

associated with an increased likelihood of a negative overall assessment.

Next, instead of using negatively reviewed technologies as controls in specification (3), we include the 22 negatively reviewed operational characteristics as controls. A negative operational characteristic $I(NegativeOper)_{i,j,t}$ equals 1 if the operational characteristic j was negatively reviewed, and equals 0 otherwise. It is sensible that negative operational characteristics would contribute to negative overall qualitative assessment. If operational characteristics are correlated with the presence of the technologies, our estimates of the technologies may be biased. We examine whether our technology estimates are robust by controlling for the potential confounders.

Columns (2a) and (2b) of Table 8 report the estimated results of the revised specification. Conditional on the negatively reviewed operational characteristics, the estimates of fuel-saving technologies shown in column (2a) are consistent with the results of column (1a); the estimates are fairly small, and all of the statistically significant coefficients are negative. That is, the presence of turbocharged, cylinder deactivation, and diesel, are associated with a reduction in the likelihood of a hidden cost. The results continue not to find the technologies themselves associated with negative overall evaluations.

Column (2b) of Table 8 reports that a negative overall review, conditional on the presence of the technologies, has a positive association with ten negatively reviewed operational characteristics: steering feel, general drivability, general handling, acceleration feel, acceleration capability, brake feel, powertrain noise, powertrain vibration, ride comfort, and fuel economy. The results not only suggest that operational characteristics in the study are the major factors in a reviewer's overall assessment, but suggest a potential explanation for the association between negatively reviewed technologies and the negative overall rating shown in column (1b). In particular, these results are consistent with a scenario that a negatively reviewed technology leads to a negatively reviewed operational characteristic, which in turn leads to a negative overall assessment. Subsection 0 and Table 7 indicate that negatively reviewed CVT and high speed automatic are positively correlated with several negatively evaluated operational characteristics related to handling, acceleration, braking, noise, ride comfort, and fuel economy. Negatively reviewed technologies may be associated with negative overall assessment through their associated negative operational impacts.

In sum, we do not find evidence that the presence of technologies is associated with negative overall assessments of a vehicle's quality. Rather, it may be that the inclusion of the technologies in some vehicle models affects overall quality via their effects on operational characteristics. With little evidence that the technologies by themselves are associated with negatively evaluated characteristics, and somewhat stronger suggestions that quality of implementation, rather than technologies themselves, affect vehicle characteristics, it appears that hidden costs are not an inevitable effect of fuel-saving technologies.

6. Limitations

The limitations of this analysis are the same as those in Helfand et al. (2016). First, this study relies on opinions of professional auto reviewers rather than vehicle buyers. We suspect that auto reviewers are more likely to notice negative vehicle characteristics and operational impacts and better able to make comparisons across vehicles than the general vehicle buyers. If so, this study may overestimate negative impacts. Second, our analysis is short-run in that we do not capture longer-term issues, such as reliability or maintenance, which are not experienced by auto reviewers. Third, vehicle models that have undergone a significant redesign may be more likely to be selected to be reviewed. If redesigned vehicles are more likely to adopt new fuel-saving technologies, our data may over-represent the presence of new technologies in MY 2014 and MY 2015. Fourth, as discussed elsewhere in this paper, our data are not sufficient to consider our results causal. Finally, and perhaps most importantly, this study relies on the assumption that auto reviews contain useful information and are not systematically biased. The fact that results are very similar between MY 2014 and MY 2015 suggests that reviewers are, at a minimum, consistent rather than random in their evaluations.

Also, it is important to note that, for some rarely mentioned technologies with small sample sizes in our data, this paper cannot answer questions about their relationships with operational characteristics. Nevertheless, we demonstrate that the use of the maximum of robust and conventional standard errors can affect conclusions drawn from the analysis, and may be especially influential for interpretation when sample sizes are small.

7. Conclusion

Energy and transportation policies have been enacted to improve vehicle fuel economy and reduce vehicle GHG emissions in many countries, including the U.S. As a variety of fuel-saving technologies have been implemented under the standards, understanding the potential hidden costs and benefits due to adoption of fuel-saving technologies contributes to understanding the full impacts of these policies.

In this paper, using professional auto reviewers' evaluations of MY 2014 and 2015 vehicles, we find that fewer than 20% of evaluations of fuel-saving technologies in individual vehicle models were negative. We then estimate the relationships of a variety of fuel-saving technologies to operational characteristics. Our results, which serve to check and validate the findings of Helfand et al. (2016), suggest that it is possible to implement these technologies without imposing hidden costs. In addition, they suggest that problems with implementation in some vehicle models, rather than something inherent in the technologies, may contribute to occasional negative operational impacts.

This paper also examines the association of the technologies with auto reviewers' overall assessments of vehicle quality. The results similarly do not provide evidence that the presence of the technologies leads inherently to negative overall ratings, but rather that negatively reviewed technologies in some vehicle models are associated with negative overall ratings. Further, the overall assessment of vehicle quality is more strongly associated with reviewers' evaluation of vehicle characteristics than the presence of the technologies. Our results suggest the importance of operational characteristics for vehicles' overall assessment.

Thus, based on MY 2014 and MY 2015 vehicles, fuel-saving technologies appear to have been adopted without significant tradeoffs for other operational characteristics. Rather than technologies themselves, the quality of implementation of the technologies in some vehicles is more likely to be associated with the quality of operational characteristics and the overall assessment. If problems arise due to implementations, rather than the inherent natures of the technologies, then it appears that automakers have the ability to mitigate any problems arising with these fuel-saving technologies.

This paper explores whether negative operational characteristics were associated with fuel-saving technology from the perspective of vehicle consumers who operate the vehicles. On the production side, successful deployment of fuel-saving technology involves costs, strategies, and the diligence of the manufacturer. This analysis suggests that manufacturers have, for the most part, risen to this challenge.

8. Declarations of interest

None.

Acknowledgements

RTI International conducted the content analysis under EPA contract EP-C-11-045, WA 3-01 and 4-08. We thank Michael McWilliams, participants at the 2017 Society for Benefit-Cost Analysis annual conference, and commenters at the EPA's Office of Transportation and Air Quality for helpful comments.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.trd.2018.08.009.

References

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? J. Econ. Perspect. 26 (1), 3–28.

Angrist, J.D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.

Autotrader, 2015. < http://www.autotrader.com/car-reviews/2015-lexus-ct-200h-new-car-review-238542# > (accessed 04/20/2018).

Caranddriver, 2014. < https://www.caranddriver.com/reviews/2014-mitsubishi-mirage-es-test-review > (accessed 04/20/2018).

Gale, Z., 2015. 2015 Honda CR-V Touring and Long-Term Update 1. < http://www.motortrend.com/cars/honda/cr-v/2015/2015-honda-cr-v-touring-awd-review-update-1/ > (accessed 08/15/2017).

Gillingham, K., Palmer, K., 2014. Bridging the energy efficiency gap: policy insights from economic theory and empirical evidence. Rev. Environ. Econ. Policy 8, 18–38.

Helfand, G., McWilliams, M., Bolon, K., Reichle, L., Sha, M., Smith, A., Beach, R., 2016. Searching for hidden costs: a technology-based approach to the energy efficiency gap in light-duty vehicles. Energy Policy 98, 590–606.

Klier, T., Linn, J., 2012. New-vehicle characteristics and the cost of the corporate average fuel economy standard. RAND J. Econ. 43 (1), 186–213.

Klier, T., Linn, J., 2016. The effect of vehicle fuel economy on technology adoption. J. Public Econ. 133, 41–63.

Knittel, C., 2011. Automobiles on steroids: product attribute trade-offs and technological progress in the automobile sector. Am. Econ. Rev. 101 (7), 3368–3399.

Krippendorff, K., 2013. Content Analysis: An Introduction to Its Methodology, third ed. SAGE Publications Inc, Thousand Oaks, CA.

MacKenzie, D., Heywood, J., 2015. Quantifying efficiency technology improvements in U.S. cars from 1975–2009. Appl. Energy 157, 918–928.

Motortrend, 2014. < http://www.motortrend.com/cars/nissan/rogue/2014/2014-nissan-rogue-first-drive/ > (accessed 04/20/2018).

Sha, M., Beach, R., 2015. Content Analysis of Professional Automotive Reviews. RTI International final report prepared for U.S. Environmental Protection Agency. Work Assignment 3-01, EPA Contract EP-C-11-045.

Sha, M., Smith, A., Beach, R., 2016. Content Analysis of Professional Automotive Reviews. RTI International final report prepared for U.S. Environmental Protection Agency. Work Assignment 4-08, EPA Contract EP-C-11-045.

U.S. Department of Energy, U.S. Environmental Protection Agency, 2014. Model Year 2014 Fuel Economy Guide. < https://www.fueleconomy.gov/feg/download.shtml > (accessed 07/06/2016).

U.S. Department of Energy, U.S. Environmental Protection Agency, 2015. Model Year 2015 Fuel Economy Guide. < https://www.fueleconomy.gov/feg/download.shtml > (accessed 07/06/2016).

U.S. Environmental Protection Agency, 2016a. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2014, EPA 430-R-16-002, < https://www.epa.gov/sites/production/files/2016-04/documents/us-ghg-inventory-2016-main-text.pdf > , accessed 08/10/2017.

U.S. Environmental Protection Agency, 2016b. Light-Duty Automotive Technology, Carbon Dioxide Emissions, and Fuel Economy Trends: 1975 through 2016. EPA-420-R-16-010, < https://www.epa.gov/sites/production/files/2016-11/documents/420r16010.pdf > (accessed 08/10/2017).

U.S. Environmental Protection Agency, U.S. Department of Transportation, 2010. Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule. Fed. Reg. 75, 25324–25728.

U.S. Environmental Protection Agency, U.S. Department of Transportation, 2012. 2017 and Later Model Year Light-Duty Vehicle Greenhouse Gas Emission and Corporate Average Fuel Economy Standards; Final Rule. Fed. Reg. 77, 62623–63200.

U.S. Environmental Protection Agency, U.S. Department of Transportation, California Air Resources Board (CARB), 2016. Draft Technical Assessment Report: Midterm Evaluation of Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards for Model Years 2022-2025. EPA-420-D-16-900, < https://nepis.epa.gov/Exe/ZyPDF.cgi/P100OXEO.PDF?Dockey=P100OXEO.PDF > (accessed 08/10/2017).