EI @ Haas WP 212R

Car Notches: Strategic Automaker Responses to Fuel Economy Policy

James Sallee and Joel Slemrod
Revised December 2012

Revised version published in
Journal of Public Economics
96(11-12), 981-999, 2012.

Car Notches:

Strategic Automaker Responses to Fuel Economy Policy*

James M. Sallee[†] Joel Slemrod[‡]

University of Chicago University of Michigan

First Draft: November 10, 2009

This Draft: December 2, 2010

Abstract

Notches — where small changes in behavior lead to large changes in a tax or subsidy — figure prominently in many policies, but have been rarely examined by economists. In this paper, we analyze a class of notches associated with policies aimed at improving vehicle fuel economy. We provide several pieces of evidence showing that automakers respond to notches in fuel economy policy by precisely manipulating fuel economy ratings so as to just qualify for more favorable treatment. We then describe the welfare consequences of this behavior and derive a welfare summary statistic applicable to many contexts.

*The authors would like to thank for helpful comments Jeff Alson, Raj Chetty, Don Fullerton, David Good, Mark Jacobsen, Tom Kenney, Chris Knittel, Matt Kotchen, Robert Mull and seminar participants at Berkeley, Chicago, Cornell, Dartmouth, Hebrew University, Michigan, MIT, the National Tax Association, NBER, University of Illinois at Chicago, and Yale. David Cashin, Patrick Giamario, Matthew Johnson, Katherine Li and Matej Mavricek provided excellent research assistance. Financial support from the California Energy Commission is gratefully acknowledged.

[†]James M. Sallee, The Harris School, University of Chicago. Email: sallee@uchicago.edu; Web: http:/home.uchicago.edu/∼sallee.

[‡]Ross School of Business and Department of Economics, University of Michigan. Email: jslemrod@umich.edu.

1 Introduction

Notches — where small changes in behavior lead to large changes in tax liability or the amount of a subsidy — figure prominently in many policies. These notches imply large, capriciously varying, local incentives to make small changes in behavior for relatively large private, but not social, rewards. Such behavioral responses erode the intended welfare benefits of policies, whose notch features are presumably justified by the increased salience and administrative convenience of policies that appear as step functions rather than smooth, continuous schedules.

In this paper we investigate these issues by focusing on notches in policies intended to encourage the production and use of fuel-efficient vehicles, what we call *car notches*. Key aspects of U.S. and Canadian policy toward motor vehicle fuel economy feature such notches. For example, under the U.S. Gas Guzzler Tax, a car with a 14.5 miles-per-gallon (MPG) rating is subject to a $4,500 tax, while a car with a 14.4 MPG rating (and as low as 13.5) is subject to a $5,400 tax, so that a tax increase of $900 is triggered by a decrease of just 0.1 MPG. Under the Canadian EcoAuto rebate program, cars that consume less than 5.5 liters of gasoline per 100 kilometers (L/100km) qualify for a $2,000 rebate, but a vehicle that consumers 5.6 L/100km (up to 6.0) receives just a $1,000 rebate.

Policy notches have a bad reputation among economists, for the reason mentioned already. Only Blinder and Rosen (1985) have risen to their defense. In the context of encouraging the consumption of a socially desirable good, they show via simulation that when general non-linear Pigouvian subsidies cannot be used in a world with multiple heterogeneous individuals, a single-notch program can improve the ratio of induced incremental consumption to either revenue cost, welfare, or revenue cost plus welfare cost, as compared to a linear subsidy. Their intuition is that, "by targeting the subsidy to those whose tastes for the favored commodity are relatively insensitive, the notch subsidy does not "waste" money on those whose consumption is not stimulated much" (p. 742). The notch is a non-linear subsidy that can, depending on the distribution of tastes, economize on the revenue loss from subsidizing inframarginal consumption, but it is an inefficient tool when used as a simple approximation a smooth schedule.

In spite of the folk wisdom that they are sub-optimal, policy notches are ubiquitous. Tax notches include the U.S. Saver's Credit, which provides a tax credit equal to a percentage of contributions to retirement savings accounts, where the credit rate is a notch function of adjusted gross income (Ramnath 2009), and the U.K. Family Credit, which applies only to families that have one adult working 16 or more hours per week (Blundell 2000). Many social programs have eligibility notches in the form of age requirements or means tests. Notches in time (a policy change takes effect on a specific date) and space (a policy changes at the border of a county, state or country) are present in most policies.

As Slemrod (2010) discusses, notches may be justified by administrative simplicity or enhanced salience. However, policy notches also induce actors to change their behavior just enough to be situated on the beneficial side of a notch. In the case of car notches, a vehicle manufacturer may have an incentive to re-engineer its cars so as to just qualify for a more advantageous policy category. Notched policies may also trigger the introduction of qualitatively new products, what Kleven and Slemrod (2009) call *tax-driven product innovation*. They note that in Indonesia, the preferential tax treatment of motorcycles relative to autos led to the creation of a new type of motorcycle with three wheels and long benches at the back seating up to eight passengers—car-like but not so car-like as to be taxed as a car. When Chile imposed much higher taxes on cars than on panel trucks, manufacturers soon offered a redesigned panel truck that featured glass windows instead of panels and upholstered seats in the back.[1]

Although presumably the fuel economy policies of the U.S. and Canada are motivated by an externality argument, so that some re-engineering is part of an anticipated and desirable response, the lumpy nature of the responses to policy notches is, *ceteris paribus*, an inefficient way to reduce the external costs of fuel consumption. In this paper we study the behavioral responses to the U.S. Gas Guzzler Tax and Canadian feebate programs and examine the welfare consequence of these programs.[2]

We also address another notch-like aspect of fuel economy policy—publicly disclosed and

[1] These examples are drawn from Harberger (1995).

[2] Many countries have notched automobile policies. For example, Sweden offers rebates up to €4000 when purchasing a hybrid car with CO_2 emissions below 120 grams per kilometer. Malaysia has notched registration and fuel subsidies that depend on engine displacement. In the U.K., the annual vehicle tax is a notched schedule determined by CO_2 emissions. Before 2001, it was a notched schedule in engine size, which may explain why there are so many cars with 1399cc engines.

highly visible fuel economy ratings that are designed to provide information to prospective vehicle purchasers. In the U.S., regulations require that automakers disclose fuel economy ratings to consumers and that these ratings be reported as *integers*. This means that, just as for the Gas Guzzler Taxes and feebate system, an underlying continuous fuel economy measure is transformed into coarser categories by policy. To the extent that consumer demand depends on the publicized ratings, and not on the underlying tests—a plausible assumption, given the difficulty of obtaining and interpreting these underlying test data—manufacturers have an incentive to re-engineer vehicles to achieve a higher integer rating.

Tax economists have recently taken interest in the study of *kinks*—points where a policy causes a discrete change in the *slope* of a tax, arguing that the extent of bunching can identify structural parameters of utility functions.[3] Often, empirical estimates have shown a more muted behavioral response to kinks than would be suggested by theory. The incentives surrounding notches are frequently much starker, suggesting that notches may prove more useful than kinks in uncovering behavioral parameters. A systematic study of the difference between kinks and notches may also shed light on phenomena related to some form of bounded rationality — if agents respond to notches but not kinks, a leading explanation might be salience.[4]

We begin in Section 2 by describing U.S. and Canadian fuel economy policies and the notches they create. In Section 3 we describe our data. Our empirical analysis begins in Section 4. There, we first show histograms of the distribution of fuel economy ratings for vehicles subject to the U.S. Gas Guzzler Tax. We find evidence of local fuel economy response, as evidenced by a statistically significant number of "extra" vehicles with fuel economy ratings just on the tax-preferred side of notches. We also show that vehicles with higher sales volume are more likely to lie on the tax-favorable side of a notch, and that there is significant bunching above the top Gas Guzzler Tax notch. We also show that the data pass falsification tests. Finally, we show evidence that automakers strategically responded to the introduction of the Canadian feebate program by modifying vehicles close to tax notches. Taken together, this evidence strongly

[3]For example, see Saez (2010).

[4]Chetty, Friedman, Olsen and Pistaferri (2009) develop an explanation of the observed response to kinked budget sets based on optimization frictions.

supports our hypothesis that automakers do respond to local notch incentives by strategically altering fuel economy ratings.

In Section 5 we turn to a welfare analysis of notches. First, we show that, under some simplifying assumptions there is a simple statistic, what we call the *average effective tax rate* around a notch, that summarizes how local manipulation distorts the intended effects of a corrective tax. Second, we demonstrate that this value, in conjunction with *ex post* aggregate data, determines a measure of the local welfare cost of using a notched policy. Third, we calculate this for the Gas Guzzler Tax data, concluding that the welfare benefits from local manipulation are negative.

In Section 6 we provide evidence that automakers also manipulate vehicle fuel economy in response to presentation notches created by rounding rules in fuel economy label regulations. This not only provides additional evidence of automaker response to notches, but it also implies that consumers value fuel economy. In Section 7 we relate the degree of bunching around Gas Guzzler Tax notches and fuel economy label notches to estimate this consumer valuation. We first show that there is greater manipulation around notches of greater tax value, and then relate the amount of bunching around the Gas Guzzler Tax to the amount around fuel economy labels in order to infer the consumer valuation of fuel economy. Section 8 concludes.

2 Fuel Economy Policy and Presentation Notches

2.1 The U.S. Gas Guzzler Tax

When in 1978 the U.S. introduced the Corporate Average Fuel Economy (CAFE) program, it also enacted the Gas Guzzler Tax, which penalizes cars with low fuel economy. The amount of the tax is a notched schedule in fuel economy, so vehicles with very small ratings differences may be subject to discretely different taxes. The tax was phased in between 1980 and 1991, but the schedule has not changed since. However, because the tax is not adjusted for inflation, the real value has eroded. Table 1 shows the schedule over time.

Light trucks, a designation that includes pickup trucks, sport-utility vehicles and vans,

Table 1: Gas Guzzler Tax Rates Over Time (Dollars Per Car)

Vehicle fuel economy (MPG)	1980	1981	1982	1983	1984	1985	1986-90	1991+
Over 22.5	0	0	0	0	0	0	0	0
22.0–22.4	0	0	0	0	0	0	500	1,000
21.5–21.9	0	0	0	0	0	0	500	1,000
21.0–21.4	0	0	0	0	0	0	650	1,300
20.5–20.9	0	0	0	0	0	500	650	1,300
20.0–20.4	0	0	0	0	0	500	850	1,700
19.5–19.9	0	0	0	0	0	600	850	1,700
19.0–19.4	0	0	0	0	450	600	1,050	2,100
18.5–18.9	0	0	0	350	450	800	1,050	2,100
18.0–18.4	0	0	200	350	600	800	1,300	2,600
17.5–17.9	0	0	200	500	600	1,000	1,300	2,600
17.0–17.4	0	0	350	500	750	1,000	1,500	3,000
16.5–16.9	0	200	350	650	750	1,200	1,500	3,000
16.0–16.4	0	200	450	650	950	1,200	1,850	3,700
15.5–15.9	0	350	450	800	950	1,500	1,850	3,700
15.0–15.4	0	350	600	800	1,150	1,500	2,250	4,500
14.5–14.9	200	450	600	1,000	1,150	1,800	2,250	4,500
14.0–14.4	200	450	750	1,000	1,450	1,800	2,700	5,400
13.5–13.9	300	550	750	1,250	1,450	2,200	2,700	5,400
13.0–13.4	300	550	950	1,250	1,750	2,200	3,200	6,400
12.5–12.9	550	650	950	1,550	1,750	2,650	3,200	6,400
Under 12.4	550	650	1,200	1,550	2,150	2,650	3,850	7,700

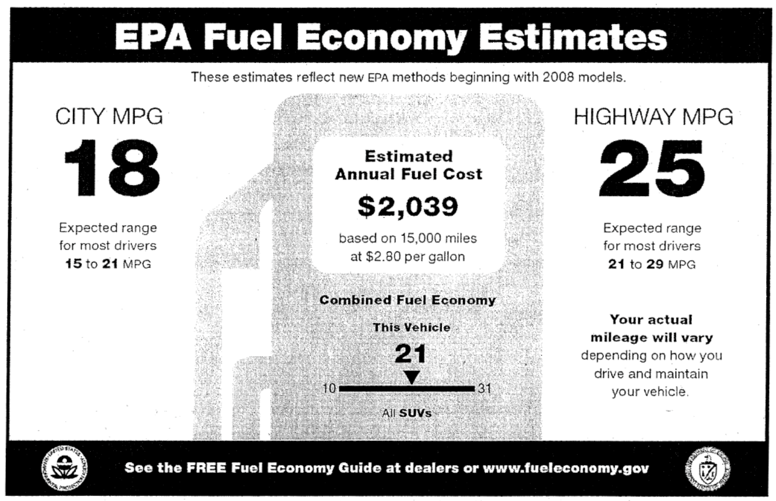
Source: Internal Revenue Service, Form 6197. All values are nominal.

were exempted from the tax from its inception, originally with the intention of not penalizing vehicles used for farming and commercial purposes. Because the Gas Guzzler Tax applies only to vehicles with low fuel economy and does not apply to light trucks, a small fraction of the market is subject to the tax. Affected vehicles tend to be high-priced, high-performance cars with relatively low sales volumes. In 2008, 77 (out of 1,248) vehicle configurations – a unique engine (including cylinders and displacement) and transmission – were subject to the tax, which raised about $172 million in revenue.[5] The tax is remitted by manufacturers, but it is visible to consumers because it appears as a separate item on the sticker price.

The fuel economy ratings used to determine tax liability are based on fuel economy tests specified by the Environmental Protection Agency (EPA). Automakers "drive" a test vehicle through a specified course on a dynamometer (essentially a treadmill for cars), during which the vehicle's exhaust emissions are captured. The amount of fuel consumed during the trial is

[5]Source: Internal Revenue Service, Statistics of Income, Historical Table 20.

Figure 1: A Sample Fuel Economy Label



Source: Environmental Protection Agency.

determined based on the quantity of several gases captured from the exhaust. Two different "courses" are used to generate separate ratings for a city and highway test. The highway and city ratings are harmonically averaged to create a single Gas Guzzler Tax rating. These same tests are also used to determine fuel economy label ratings, CAFE ratings, and for emissions regulations. Each rating involves a slightly different transformation of the underlying test results, so that a vehicle may be near a notch for one rating system but not for others.

2.2 Fuel Economy Label Ratings

Every new vehicle sold in the United States is required to display a label that details the vehicle's MSRP, and since 1978 this label must also include the vehicle's official EPA highway and city fuel economy ratings. The font size each label item is mandated by law, and the city and highway fuel economy ratings must be set in the largest font, making these by far the most prominent numbers on the label. In a much smaller font, the label also displays the combined rating, a graphic that compares the vehicle to others in its class, and an estimate of the annual cost of gasoline. Figure 1 is an example of the current fuel economy label in the United States.

The city and highway ratings are integers, which are determined by rounding off the underlying fuel economy estimate derived from the test procedure. This rounding creates what we call a *presentation notch* — where a marginal difference in an underlying characteristic creates a discrete change in the information transmitted in the marketplace. A vehicle with a highway fuel economy rating of 29.49 will be listed as 29 on the label, whereas a vehicle with a rating of 29.50 will be listed as 30.

If consumers value fuel economy, and if they use the official EPA ratings as a source of information, then firms may undertake costly adjustment procedures to increase the fuel economy rating as displayed on the labels, just as they would respond to tax notches.[6] The label ratings are based on the same tests used to derive the Gas Guzzler Tax rating. Prior to 1986, the reported label rating was simply the integer nearest the value resulting from the test procedure.[7] Starting in 1986, the EPA modified the procedure to adjust for "in-use shortfall", in response to the fact that consumers consistently reported lower average actual fuel economy than suggested by the labels. After attempting to measure the discrepancy, the EPA decided to adjust the test numbers by simply multiplying the test output by a fixed factor—0.9 for the city and 0.78 for the highway test. The product is then rounded to the nearest integer for the label.[8] Automakers do have the right to adjust the label ratings downwards if they wish, and a very small percentage of ratings do reflect a downward adjustment, so that the test procedure indicates a higher value than appears on the label in practice. The automakers may do this to avoid consumer displeasure if actual fuel economy experiences fall short of their expectations.

Fuel economy labels are assigned according to basic engine and transmission. Thus, separate label ratings are not reported for vehicles that share a basic engine and transmission but have different vehicle weights. Testing is required, however, for each vehicle with a different weight, and the test results are combined via a sales-weighted harmonic average.[9]

[6]The unrounded figures are public information (indeed, we make use of these data in subsequent analysis), and so in principle consumers could also obtain and consider these unrounded numbers. To do so, however, they would have to download the publicly available fuel economy data files; the unrounded numbers are not included in the fuel economy guide that is available at dealerships and on the EPA's website. Even with these files in hand, a car shopper would need to know how to adjust for a factor called "in-use shortfall" in order to convert the unrounded numbers. We think this is unlikely, but if consumers do find the unrounded information, we would not expect automakers to respond to label notches, which we test directly below.

[7]The EPA uses ASTM International rounding, which rounds a value ending in exactly 0.50 to the nearest even integer.

[8]Starting in 2008, the EPA instituted a new testing procedure, which is designed to improve the accuracy of label ratings. In this paper, we investigate only models that were tested during the pre-2008 regime.

[9]Tests are not performed on every model separately if models share the same basic engine, transmission and

Table 2: Rebate and Tax Thresholds in the Canadian Feebate Program

Fuel Economy		**Rebate**			**Tax**
L/100km	MPG Equivalent	Cars	Light Trucks	Flexible-Fuel Vehicles	Cars, Vans and SUVs
5.5 or less	42.8 or more	$2,000			
5.6 - 6.0	42.0 - 39.2	1,500			
6.1 - 6.5	38.6 - 36.2	1,000			
7.3 or less	32.2 or more		$2,000		
7.4 - 7.8	31.8 - 30.2		1,500		
7.9 - 8.3	29.8 - 28.3		1,000		
13.0 or less	18.1 or more			$1,000	
13.0 - 13.9	18.1 - 17.0				$1,000
14.0 - 14.9	16.8 - 15.8				2,000
15.0 - 15.9	15.7 - 14.8				3,000
16.0 and over	14.7 or less				4,000

Note: The feebate program measures fuel economy in L/100km. To facilitate comparison with the U.S. Gas Guzzler Tax, we show the miles-per-gallon equivalent.

2.3 The Canadian Feebate Program

In March 2007, the Canadian government introduced two new notched fuel economy programs. The first was called the Green Levy, which taxed particularly fuel-inefficient vehicles. As with the Gas Guzzler Tax, pickup trucks were exempted, but sport-utility vehicles and vans were subject to the tax. As shown in Table 2, the maximum tax was $4,000. In the 2008 model year, 156 (out of 1071) distinct vehicle configurations were taxed. Because many of the vehicles subject to the tax are low-volume, high-performance vehicles, the taxed vehicles comprise a larger fraction of models than their share of the total sales volume, which was 2% in 2008.

The second policy, called the EcoAuto rebate program, was introduced simultaneously. It provided rebates for vehicles with particularly good fuel economy. Unlike the Green Levy, the EcoAuto program was designed to be a temporary measure, set to expire after two years. As a result, the program was in effect from March 20, 2007 to December 31, 2008, and only vehicles in model years 2006, 2007 and 2008 were eligible to receive the rebate. In the 2008 model year, only 32 vehicle configurations qualified out of 1071 vehicle configurations, but these vehicles enjoyed an 8% market share. Together the rebate and the tax comprise a version of what is

weight. For example, a Mercury Mountaineer is identical to a Ford Escape in engine, transmission and weight. As a result, Ford Motor Company would test only one of the two models.

sometimes called a feebate — a set of taxes and rebates that together act to encourage the purchase of more fuel-efficient vehicles. The rebate program cost $191.2 million on the program over two years, but the magnitudes of the two programs were expected to be roughly offsetting, making the combined policy revenue-neutral.

In Canada, fuel economy ratings are measured as liters of gasoline consumed per 100 kilometers (L/100km), which differs from MPG both in units and in inverting the ratio of fuel to distance traveled. The fuel economy ratings used in Canada are based on test procedures that are nearly identical to the ones used in the United States. Cars that consume less than 6.5 L/100km were eligible for a subsidy of at least $1,000. Light trucks, including pickups, SUVs and vans, were subsidized if they consume less than 8.3 L/100km. Flexible-fuel vehicles — vehicles capable of running either on conventional gasoline or a fuel blend known as E85 that is 85% ethanol — that got 13L/100km or better were eligible for a $1,000 rebate. All vehicles in the eligible vehicle classes were subject to the tax if they consume above 13 L/100km. Starting at $1,000, the tax rose by $1,000 for each integer increase in L/100km, for a maximum of $4,000. Table 2 summarizes these policies. The current values of both the Canadian policies and the Gas Guzzler Tax are also plotted in Figure 2.

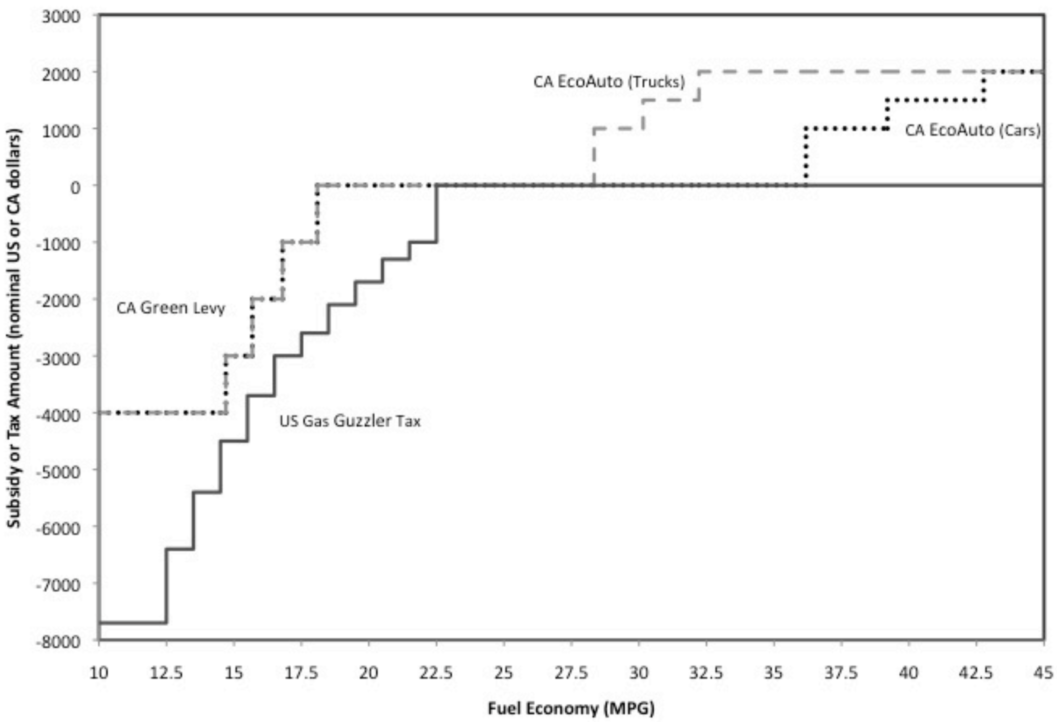
The introduction of the feebate program in March 2007 was said to be a surprise to automakers and consumers alike. This provides an opportunity to examine automaker response to notches by comparing 2007 model year vehicles, which were designed and tested before the policy was revealed, and 2008 and 2009 model year vehicles, which could have been modified in response to the policy. We take this up in section 4.5.

2.4 How Is Fuel Economy Manipulated?

Our analysis is based on the premise that automakers perform local manipulation of fuel economy ratings in order to move over a tax or presentation notch. If an automaker wishes to boost fuel economy locally around a notch, how is this done?

First, automakers may simply repeat the underlying fuel economy test if there is sufficient variability across trials. U.S. regulation requires, however, that all valid tests must be reported

Figure 2: The U.S. Gas Guzzler Tax and the Canadian Feebate Program in 2008



and averaged, limiting the efficacy of this strategy.

Second, each "model type" receives a single rating, but a single model type may involve several test vehicles of different weights, which are then averaged to create a single rating. Automaker could produce less of one configuration and more of the other to move the average rating, which determines the tax liability for all vehicles in a model type, but this is likely to be an expensive strategy.

Finally, an automaker may modify a vehicle to improve its fuel economy. Methods include "light-weighting" (substituting vehicle parts to reduce weight), engine recalibration (reprogramming the vehicle to operate in a different gear at certain speeds), use of low-friction lubricants, modifications to tires, or small aerodynamic changes such as the addition of a spoiler, side skirts, air dam reshaping, or the installation of "belly pans" that smooth air flow by covering parts underneath the vehicle.[10] It is important to distinguish these methods from what one might characterize as global design choices. The overall structure of the engine, weight, the

[10]These examples are drawn from Edmunds.com and National Research Council (2002). Aerodynamics influence fuel economy ratings because testers take vehicles on an actual driving course and "coast" the vehicle to tests its aerodynamic efficiency. The results are used to make small modifications to the dynamometer output.

shape of the car, the use of fuel saving technologies and the choice of transmission are all key determinants of fuel economy, but these decisions are made on a several year time horizon, long before vehicles are officially tested and the exact location of a vehicle vis-à-vis a notch is known.[11] In contrast, the relatively minor adjustments we listed above could potentially be adopted late in the production cycle, in response to preliminary test results.

Automakers are reluctant to publicize information about how local fuel economy adjustment might occur or say whether they respond to notches. However, our conversations with experts who have worked for automakers and officials at the EPA indicate that this type of vehicle modification does indeed take place. Anecdotal evidence from the popular press also provides support. Canadian media reported that automakers intentionally altered some vehicles' fuel economy in response to the feebate (Keenan 2007a). In 2009, when the recent cash-for-clunkers bill was passed, Nissan stated its intention to alter fuel economy for certain models to ensure that they met fuel economy eligibility requirements (Greimel 2009). Another oft-cited example concerns what is known as Computer Aided Gear Selection, sometimes called a "skip shift", which forces a manual transmission vehicle into a first-to-fourth gear shift at certain speeds. Popular consensus is that this feature is installed as a way of reducing the Gas Guzzler Tax, and kits are available that claim to disable this feature. In the remainder of the paper, we empirically test for evidence of this strategic manipulation, and this background provides a plausible interpretation of our results.

3 Data Sources

We gathered fuel economy data from several sources. For the Gas Guzzler Tax, we obtained from the Internal Revenue Service a complete list of all vehicles that were subject to the tax from the beginning of the program in 1980. These data include fuel economy ratings to a tenth of a mile-per-gallon, and are limited to the set of vehicles that were actually taxed, but are complete for all years between 1980 and 2009.

[11] Longer-horizon global choices are examined by Klier and Linn (2008), Knittel (2009) and Whitefoot, Fowlie and Skerlos (2010).

We complement this data with fuel economy ratings from the EPA, which provides unrounded city and highway test results from 1978 to 1983 and from 1999 to 2007. We use these underlying test statistics to reconstruct the Gas Guzzler Tax rating for all vehicles using the formulas published in federal regulations. The EPA data have the advantage of allowing us to calculate the Gas Guzzler Tax rating for vehicles that were not subject to the tax, but it has the disadvantage of a coverage gap. Between 1984 and 1998, the EPA data do not include the unrounded test results necessary for calculating the correct Gas Guzzler Tax rating.

These same EPA data are our source for fuel economy label ratings. We transform the unrounded city and highway test results according to the EPA's in-use shortfall adjustment factors in order to obtain the adjusted, unrounded fuel economy label ratings. In our analysis of CAFE ratings, we use official CAFE fuel economy ratings, which differ slightly from the EPA ratings used in the other sections, from the National Highway Transportation Safety Administration (NHTSA). CAFE data include sales volumes, and we match the CAFE data set to the IRS Gas Guzzler list in order to measure the sales volumes of taxed vehicles. We use this source of sales data instead of more conventional sources like Automotive News because conventional data sources do not divide the sales of a model line among the different engine configurations, so that the unit of observation is not the same across data sources. Our Canadian data come from program documents available from the Canadian Government.

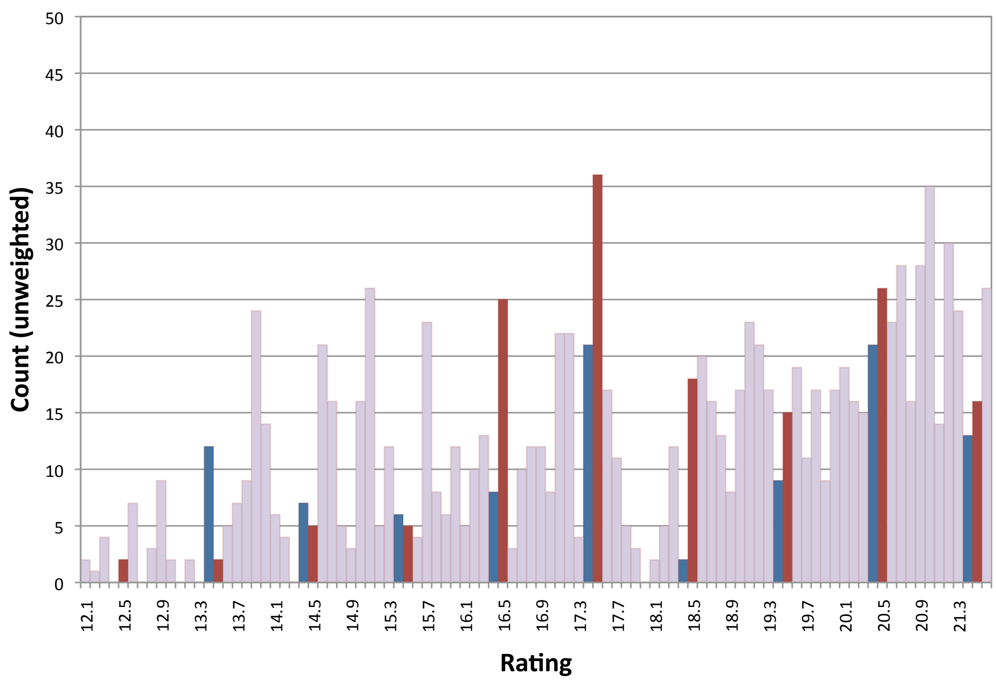
4 The Behavioral Response to Fuel Economy Notches

4.1 The Gas Guzzler Tax

If automakers respond to notches in the Gas Guzzler Tax by modifying vehicle fuel economy, then the distribution of fuel economy ratings should feature "extra" observations just on the tax-preferred side of notches. A natural first test, then, is to examine the distribution of the rating decimal points of all taxed vehicles, relative to the tax notches.

Figure 3 is a histogram of the number of models by their one-decimal Gas Guzzler Tax rating between 1991 and 2009, during which time the Gas Guzzler Tax schedule is stable and

Figure 3: Gas Guzzler Rating Distribution, Unweighted: 1991 - 2009

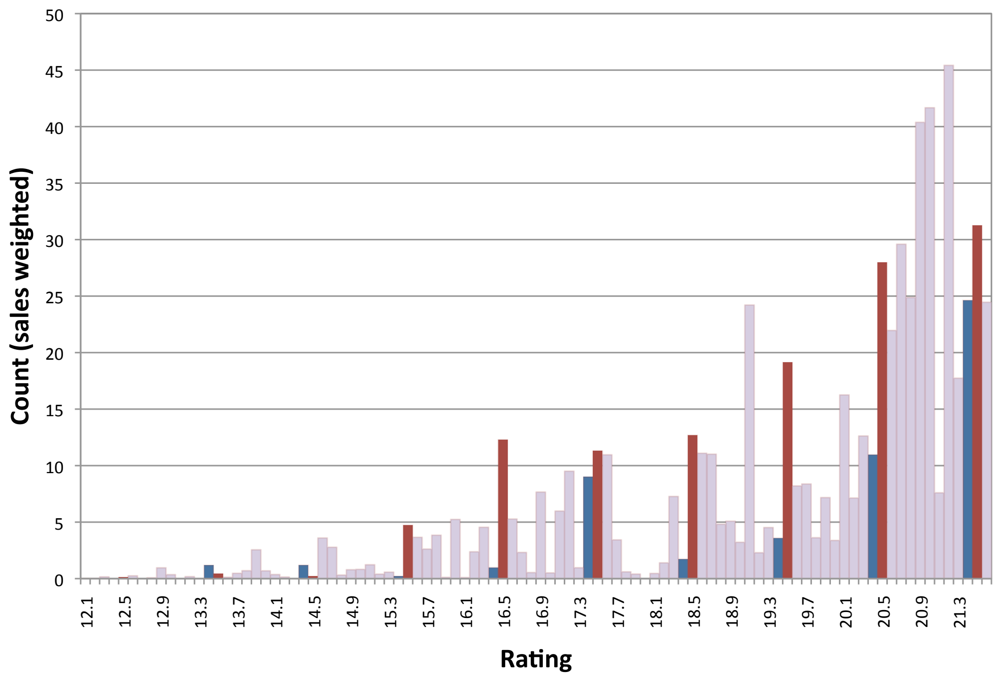


Note: IRS data, sample size is 1,221. Ratings ending in .4, all of which are just below a tax notch, are colored in blue, while ratings ending in .5 are colored in red.

notches are present at each rating ending in .5. Bars at a .4 decimal (the low-tax side of a notch) are shaded blue, and bars at a .5 rating (the high-tax side of a notch) are shaded red. Of the ten different integer values, the number at .5 exceeds the number at .4 in seven cases, sometimes by a large margin. Overall, there are 150 models at a .5 decimal and 99 at a .4 decimal. The probability that, of 249 draws, 150 or more would be drawn from a binomial distribution with equal probability is just 0.0007. If we compare the number of models at .3 or .4 to the number at .5 or .6, the story does not change: 200 just below the notch versus 295 just above.

The counterexamples to the preponderance of .5 decimals over .4 decimals are high-performance, high-price ultra-luxury automobiles with very low fuel efficiency. Manufacturers of these cars may perceive that their prospective buyers care little about a few hundred dollars because it is a small fraction of the total cost, or even perceive that a low MPG is a status symbol of high performance. These models also have relatively low sales volume, so that if modification involves a fixed cost, we would see less bunching among these vehicles. To capture this possi-

Figure 4: Gas Guzzler Rating Distribution, Sales Weighted: 1991 - 2007



Note: IRS fuel economy data and NHTSA sales data; sample size is 841. Ratings ending in .4, all of which are just below a tax notch, are colored in blue, while ratings ending in .5 are colored in red. Sample differs from figure 3 because some vehicle types are missing sales information and sales data are unavailable for 2008 and 2009.

bility, figure 4 replicates figure 3 but weights the distribution by sales volume. In this figure, the predominance of .5 decimals is even more pronounced, and the integers where .5 do not predominate feature very low sales.

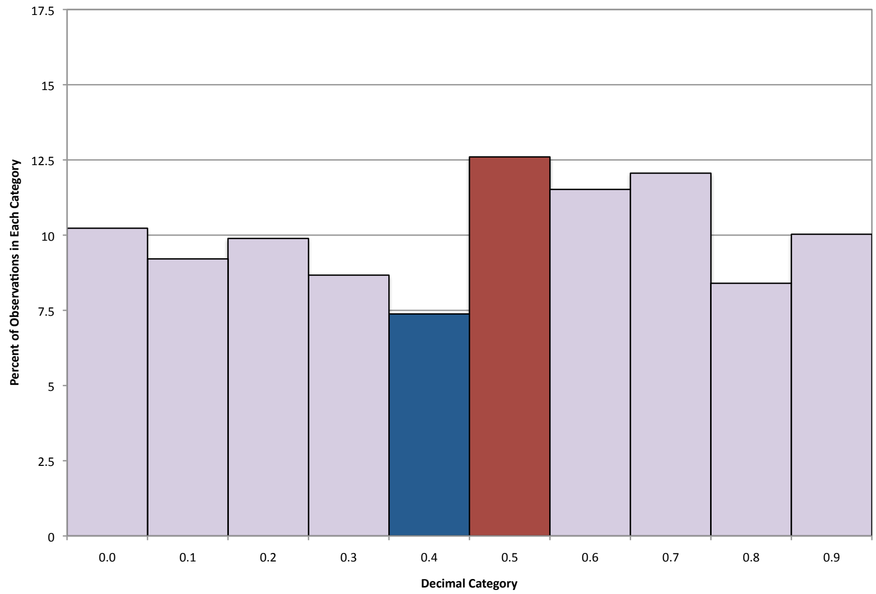
Figure 5 aggregates across integers to show a histogram of MPG of the *decimal* values for all vehicles subject to the Gas Guzzler Tax. For example, if a vehicle had a 20.5 fuel economy rating, we put that vehicle into the .5 bin. Aggregation allows us to include more data by combining different tax regimes.[12]

Absent tax incentives, we might expect this decimal distribution to be uniform. The actual distribution shows a marked departure from uniformity, with far more observations just at, or just to the right of .5.[13] This difference is highly unlikely to be due to chance. Comparing

[12]As shown in Table 1, the value of the Gas Guzzler Tax changes at each .5 in fuel economy ratings, except for 1980, 1981, 1983 and 1985, for which years we adjust the data to match the .5 notch point in the figure.

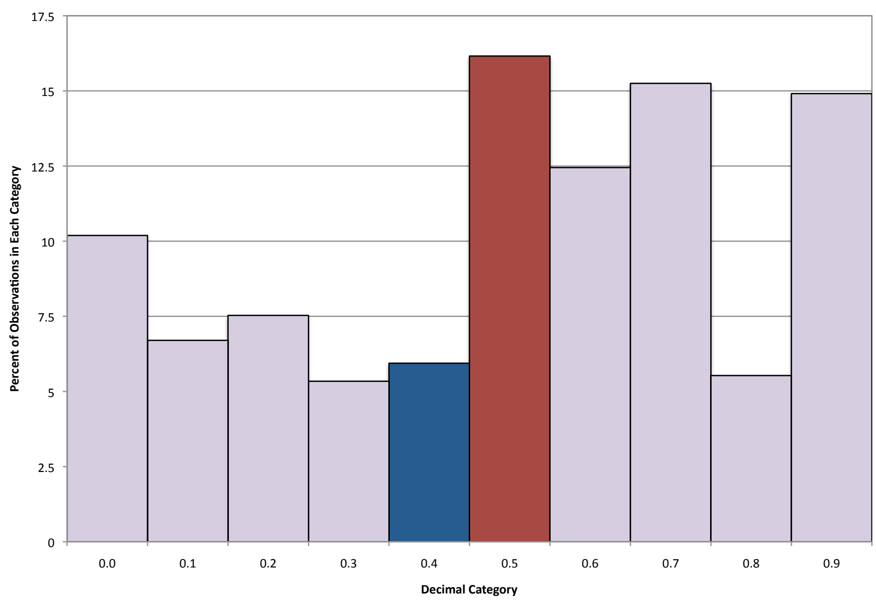
[13]Results are qualitatively similar if we restrict the sample to unique observations by dropping all vehicles with the same manufacturer, cylinders, displacement, transmission and fuel economy rating as some other vehicle, either within or across years. This restriction is intended to drop repeated observations of the same engine,

Figure 5: Gas Guzzler Decimal Distribution, Unweighted: All Vehicles, 1980 - 2009



Note: IRS data; sample size is 1,476. In several years, the notch is at whole integers (the .0 bin). For those years, we shift decimals by .5, so the notch is always represented by the .5 bin.

Figure 6: Gas Guzzler Tax Decimal Distribution, Sales Weighted: All Vehicles, 1988 - 2007



Note: IRS and NHTSA data; sample size is 945.

either the number of vehicle configurations in the .4 bin to the .5 bin, or comparing the sum of the .3 and .4 bins to the sum of the .5 and .6 bins, yields a p-value less than .0001 that they are

which may be installed on several different models.

drawn from a uniform distribution.[14] Figure 6 provides a sales-weighted histogram of ratings decimals. Here, the gap between sales around the notch are even more pronounced, though the distribution shows greater variation overall.

The statistical tests cited above are based on the assumption that, in the absence of notch responses, the preponderance of .4 and .5 decimals would be the same. This assumption may not be precisely correct if the overall fuel economy distribution has a positive slope, in which case there might be more .5 decimals for reasons unrelated to the notches. If this were driving our results, we would expect figures 5 and 6 to show a tilt across all decimals, i.e., there would be more .1 than .0, more .2 than .1, etc. We do not see this pattern.

To further dispel such concerns, we redid our statistical tests after accounting for the overall shape of the fuel economy distribution, the results of which we report in table 3. First, we estimate a polynomial through the frequency distribution in figure 3, omitting observations at the .4 and .5 decimals. We then use the predicted values from these polynomials to predict the relative number of .4 and .5 decimals that should occur, given the shape of the distribution. Combining these estimates yields the predicted probability that a vehicle would have a .5 decimal, conditional on the observation being either .4 or .5, under the null hypothesis that the polynomial predicts the relative prevalence correctly. We then use this new predicted probability to ask how likely it is that we would have observed 150 observations at .5 out of 249 that were either .4 or .5. Rather than simply do a single t-test with the adjusted probabilities, however, we bootstrap this entire procedure (starting by resampling our microdata) so as to incorporate the variance that arises from the estimation of the polynomial.

Table 3 shows that this adjustment has very little impact on the estimated probabilities. The first row of the table shows what we label the binomial model, which is our original assumption that the counterfactual probabilities of ratings ending in .4 and .5 are equal. We observe 150 out of 249 above the notch. Under the binomial model assumption, the expected number of observations above the notch is 124.5, and the standard deviation is 7.89. The second row

[14] This significance test comes from treating the observations in the restricted distribution just around the notch as a binomial distribution, with points above the notch treated as a successful trial. We calculate the p-values reported here using the normal approximation to the binomial distribution.

Table 3: Statistical Tests of Bunching Above Gas Guzzler Tax Notches

	Observed Number Above Notch (Out of 249 Within .1 MPG)	Expected Number Above Notch Under Null Hypothesis	Standard Deviation of Number Above Notch Under Null Hypothesis	p-value
Binomial Model	150	124.5	7.89	0.0007
Linear Control Function	150	125.4	7.89	0.0009
Quintic Control Function	150	124.8	7.97	0.0008

Note: The binomial model assumes that the probabilities that a vehicle has a fuel economy rating ending in .4 and .5 are the same. The linear model modifies this assumption by adjusting for the overall shape of the distribution using a linear fit, excluding observations within .1 MPG of a notch. The quintic model extends this by using a fifth order polynomial to estimate the distribution's shape. Statistics for the linear and quintic models are derived via bootstrap to account for sampling variation in the estimated polynomial.

of table 3 shows the expectation and standard deviation when only a linear control is used, and the third row shows a fifth-order polynomial. The expected number of observations above the notch and the standard deviation change only slightly. This is not surprising because the overall shape of the distribution in figure 3 does not exhibit a dramatic slope. The probability of observing 150 observations above the notch is extremely unlikely under any of the modeling alternatives, which bolsters the conclusion that the data reveal strategic responses to notches.[15]

4.2 Determinants of Bunching

If automakers respond to notches, we would expect to find more bunching around notches with greater tax values. And, if there are fixed costs in modification, we would expect more bunching among vehicles with higher sales. We test these additional hypotheses using a linear probability model, where the dependent variable is coded as 1 if an observation falls on the tax-preferred side of a notch, in a sample restricted to cars within a window on either side of notches.

Table 4 shows results using a window of .2 MPG around notches, so the sample includes vehicles whose Gas Guzzler Tax rating ends in either .3, .4, .5 or .6, and the dependent variable is coded as 1 for the .5 and .6 observations. The first four columns use data from the IRS. This

[15] An alternative methodology is to collapse the data and perform statistical tests on the aggregated cell counts, treating each fuel economy rating as the unit of observation rather than each vehicle. Regressions based on this approach are included in the appendix. They show a large amount of bunching, which is statistically significant when cell counts are sales weighted.

Table 4: The Correlation Between Bunching and Sales Volume and Tax Notch Value

Dependent variable = 1 if vehicle on tax-preferred side of notch
Sample includes vehicles with ratings ending in .3, .4, .5 and .6

	IRS Data (1980-2009)				EPA Data (1998-2007)	
	1	2	3	4	5	6
Log Sales		0.0439***		0.0435***		
		(0.0149)		(0.0161)		
Gas Guzzler Value ($100)			0.00179	-0.000548		0.0136***
			(0.00684)	(0.00679)		(0.00393)
Gas Guzzler Range Dummy					0.126***	
					(0.0328)	
Constant	0.600***	0.340***	0.589***	0.345***	0.447***	0.450***
	(0.0201)	(0.0931)	(0.0463)	(0.0463)	(0.0113)	(0.0111)
Observations	593	371	593	371	2,207	2,207
R^2	0.00001	0.023	0.0001	0.023	0.007	0.005

Note: Columns 1 through 4 use IRS data, which include only vehicles that pay the Gas Guzzler Tax. Columns 5 and 6 use EPA data. This enables the inclusion of vehicles that do not pay the Gas Guzzler Tax, including both vehicles located just above the top notch and those further away who face a zero incentive for moving over .5. Sales data are not available in the EPA data, so not all specifications are repeated with both samples.

allows us to include all years of data, from 1980 to 2009, and to include sales volume, but it restricts the sample to observations that pay the Gas Guzzler Tax. This means that we cannot include the top notch, nor can we include vehicles that are not around a tax notch as a baseline. Columns 5 and 6 use EPA data instead, which allows us to include the top notch and vehicles not near a notch as a baseline, but it restricts the years of availability and does not allow us to include sales volume controls due to data limitations.

Column 1 includes only a constant, which is equal to the fraction of observations on the tax-preferred side of notches. This estimate is statistically different from .5, which would be the counterfactual value under the null hypothesis of no strategic bunching. Column 5 provides a similar test in the EPA data, but it includes both a constant and a dummy variable coded as 1 if the vehicle is in the range of the Gas Guzzler Tax because there are also vehicles in the sample that are not near a notch. The fact that the dummy is positive and statistically significant indicates that vehicles facing a Gas Guzzler Tax notch are more likely to have a decimal of .5 or .6 than vehicles with fuel economy outside the Gas Guzzler Tax range.

Column 2 shows that vehicles with higher sales volumes are significantly more likely to be on the tax-preferred side of a notch. If automakers were not bunching strategically and the "extra" observations on the tax-preferred side of notches were due to chance, there would be no reason to expect those observations on the tax-preferred side to be higher volume.[16] If, however, there are fixed costs in manipulation, we would expect higher sales volume vehicles to be more likely to bunch. The strong correlation between sales and bunching therefore provides additional support for our hypothesis that automakers respond to notches.

The IRS data (column 3) shows a small, statistically insignificant coefficient on notch value. In the EPA data (column 6), however, there is a positive coefficient, so there is more bunching surrounding notches with higher values. The difference is due to the importance of the top notch, which has many data points and a large notch value of $1,000. Excluding data from the top notch in the EPA data produces a coefficient similar to the IRS data (not shown).

4.3 False Experiments

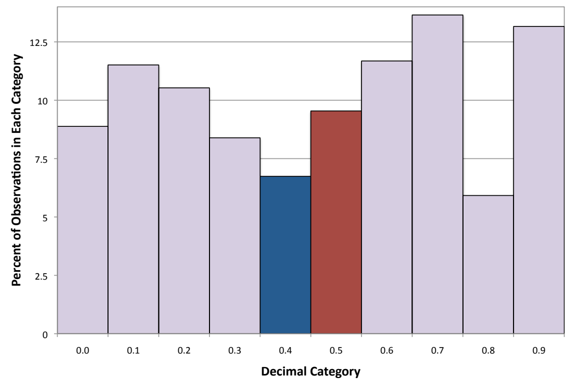
One might be concerned that the preponderance of vehicle configurations with ratings at or just above .5 might be an artifact of some unknown engineering property or other anomaly. We test this by reproducing the Gas Guzzler Tax fuel economy measure for vehicles not subject to the tax and seeing if the pattern reemerges here. This requires use of the EPA data, which limits the sample to 1999-2007. First, we reproduce the main result from figure 5 for this subsample of years in part (a) of figure 7, which shows evidence of bunching but with considerably greater noise than the main figure with the unconstrained sample.

Second, we look for bunching behavior around notch values when no tax incentive for bunching exists. Part (b) shows the distribution of fuel economy decimals for passenger cars that have ratings above the Gas Guzzler Tax threshold MPG and therefore have no incentive to bunch at or just above .5 decimals. Likewise, part (c) shows the distribution of rating decimals for light trucks in the same fuel economy range as the passenger cars in part (a). Because light

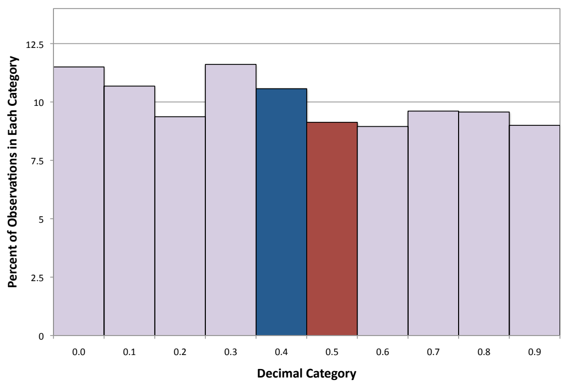
[16] Here we describe sales volume as an exogenous characteristic, but a lower Gas Guzzler Tax for a vehicle will lead to a lower price and thus increase sales, making sales an endogenous characteristic. This is true, but an extraordinarily high price elasticity would be required to explain the magnitude of the differences in sales volume for high and low tax vehicles observed in the data.

Figure 7: Gas Guzzler Rating Decimals for Several Vehicle Groups: 1999 - 2007

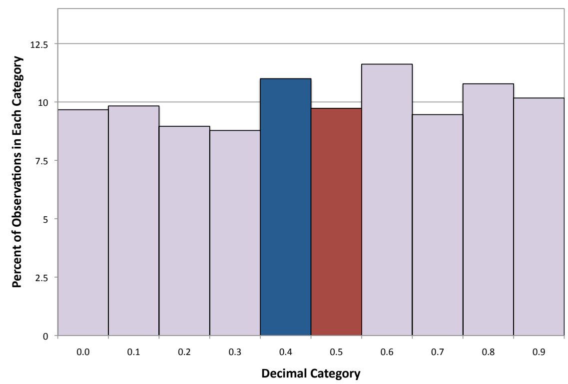
(a) Passenger Cars Subject to Tax



(b) Passenger Cars Not Subject to Tax



(c) Light Trucks with Fuel Economy in Taxed Range

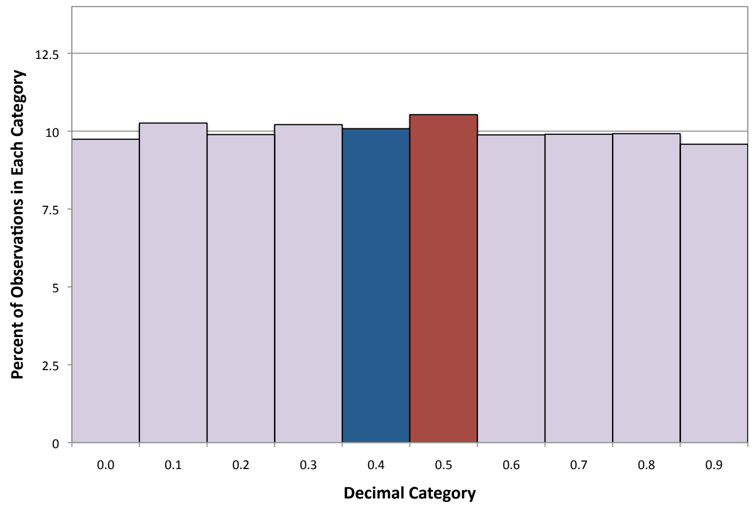


Note: Total sample sizes are (a) 608, (b) 5,422 and (c) 3,236. Vehicles in parts (b) and (c) are not subject to the tax and not expected to exhibit bunching.

trucks are not subject to the Gas Guzzler Tax, there is no incentive to bunch. The fact that neither class of vehicles exhibits bunching is further evidence that the bunching in the vehicles subject to the tax is due to a strategic response to notches.

For an additional false experiment, we examine a closely related fuel economy measure, individual vehicle CAFE ratings, which do not have a notch at .5. Each vehicle in a manufacturer's fleet is given a CAFE rating based on a weighted average of the vehicle's city and highway fuel economies. These combined ratings, which are calculated to the tenth of a MPG (e.g., 27.5 MPG), are used to calculated a sales-weighted average for all vehicles made by a given manufacturer. This sales-weighted average is then rounded to a tenth of a mile-per-gallon for

Figure 8: Distribution of CAFE Combined Rating: 1978 - 2005



Note: NHTSA data; sample size is 18,045.

use in determining compliance with CAFE. Because the individual fuel economy ratings are not rounded to integers prior to averaging, there is no incentive for manufacturers to push individual vehicle CAFE ratings above any particular decimal.[17] Figure 8 shows the combined CAFE rating decimal distribution. The ratings are roughly uniform, as expected. There are slightly more observations in the .5 bin than the .4 bin, but this difference is not statistically significant.[18] This is further evidence of our main conclusion.

4.4 Bunching Above the Top Gas Guzzler Notch

Looking only at the fuel economy ratings used in the IRS data does not reveal what is arguably the most striking example of bunching: increasing fuel economy so that it just above the threshold for any tax at all. It is not apparent because the IRS does not publish ratings for vehicles that are not subject to tax, and so their data do not reveal how many models have a fuel economy rating just over the taxable threshold, which since 1991 has been 22.5 MPG.

[17] Manipulation of the Gas Guzzler will translate into manipulation of the CAFE rating, because the two numbers are identical in early years and extremely close to each other in later years. Thus, we omit passenger cars with combined fuel economy ratings below 23, which would be subject to the Gas Guzzler Tax.

[18] A test of the difference between the .4 and .5 bins yields a one-sided p-value of .092, and a test of the difference between the .3 and .4 bins from the .5 and .6 bins yields a one-sided p-value of .400. Overall, a chi-squared test statistic of the null hypothesis that the data are distributed multinomial with equal probability on each bin cannot be rejected (p-value of .994).

In order to ascertain the amount of bunching at the top, we reconstruct the Gas Guzzler Tax rating for all vehicles using EPA data in the available years — 1978 to 1983 and 1998 to 2007. Between 1978 and 1983, the Gas Guzzler Tax was changing. As indicated above in table 1, in 1978 and 1979, there was no tax. The tax began in 1980, at which time it had a top notch of 15.0. This changed to 17.0 in 1981, and then 18.5 in 1982, and finally 19.0 in 1983. Figure 9 shows the distribution of fuel economy ratings for passenger cars in each of these six years. In each diagram, the dashed blue vertical lines indicate the location of *future* top notches. The unbroken red vertical lines indicate the effective top notch for the year shown.

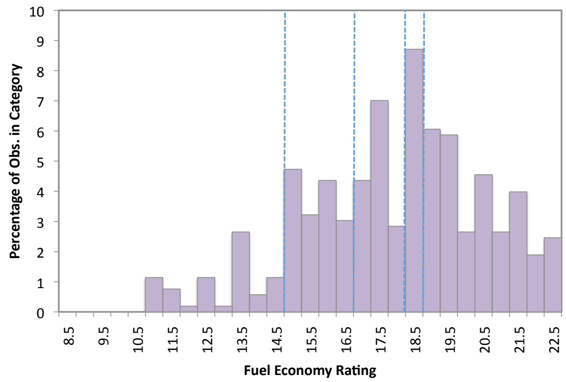
These six figures suggest a precise response to the top notch. Before the policy, a large fraction of vehicles lay to the left of the blue lines, so they would be subject to a tax in future years. When the tax is introduced for cars below 15.0 MPG in 1980, a majority of the vehicles that were previously below this level are gone. The same adjustments occur in 1981, 1982 and 1983; in each year, most of the vehicles that would have been just below the notch have moved. The entire distribution shifted rightward, not just vehicles near the notch.[19] The overall shift is likely due to CAFE, which was introduced at the same time, but it appears that the details of the distribution were driven by the location of the top Gas Guzzler Tax notch.

The EPA data are unavailable during the rest of the tax's phase-in period, but they are available for several years after the top notch of 22.5 MPG was established in 1991. Figure 10 is a histogram of the number of models of all vehicles produced from 1991 to 2007 by their CAFE fuel rating, whether they are subject to the Gas Guzzler Tax or not. The dark vertical line is drawn at the tax threshold of 22.5 MPG. The vast majority of models are above the Gas Guzzler Tax threshold, and there is a clear asymmetry in the histogram. To quantify the bunching above the top notch apparent in figure 10, we estimate an 8th-order polynomial on the distribution of passenger car Gas Guzzler Tax ratings over the available sample period from 1998 to 2007, omitting data within 1 mpg of the top notch at 22.5. We then take the predicted number of observations near the Gas Guzzler Tax top notch and compare that to the

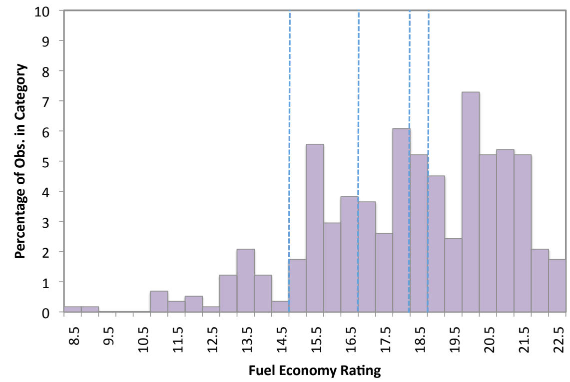
[19] This is evident in figure 9 from the reduction in the fraction of vehicles getting under 22.5 MPG. The entire distribution is not shown to preserve visibility of the relevant portion, but the heights of the histograms are preserved to be comparable.

Figure 9: Gas Guzzler Tax Rating Distribution During Tax Phase In

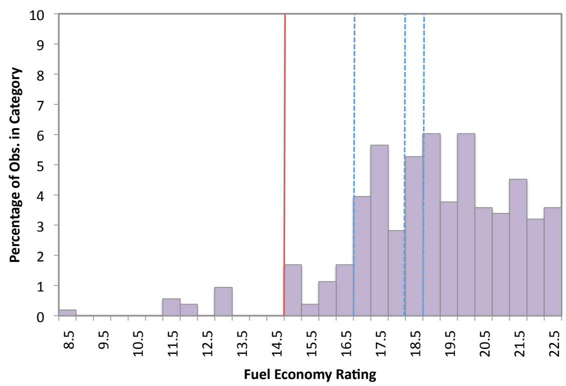
(a) 1978 (No Tax)



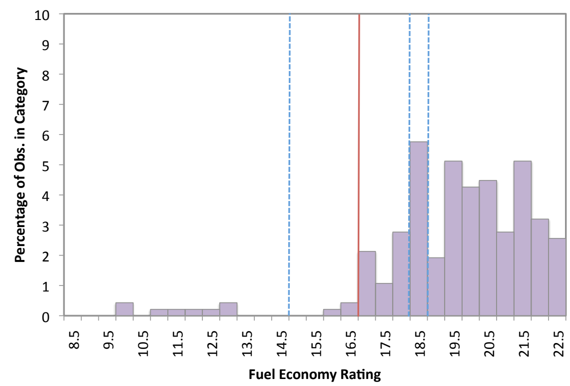
(b) 1979 (No Tax)



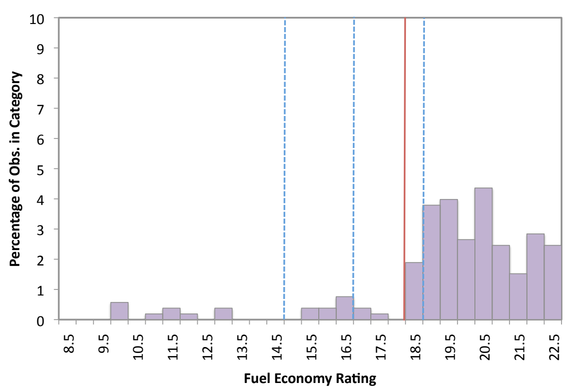
(c) 1980 (Top Notch = 15.0)



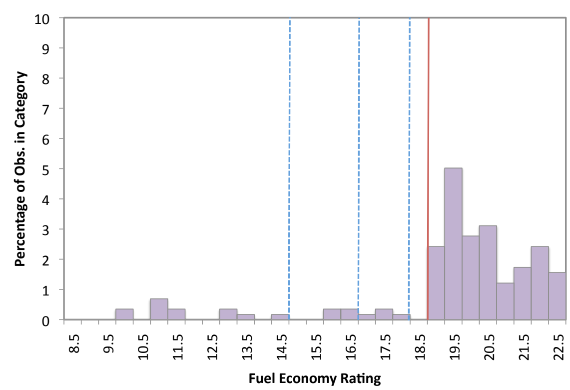
(d) 1981 (Top Notch = 17.0)



(e) 1982 (Top Notch = 18.5)



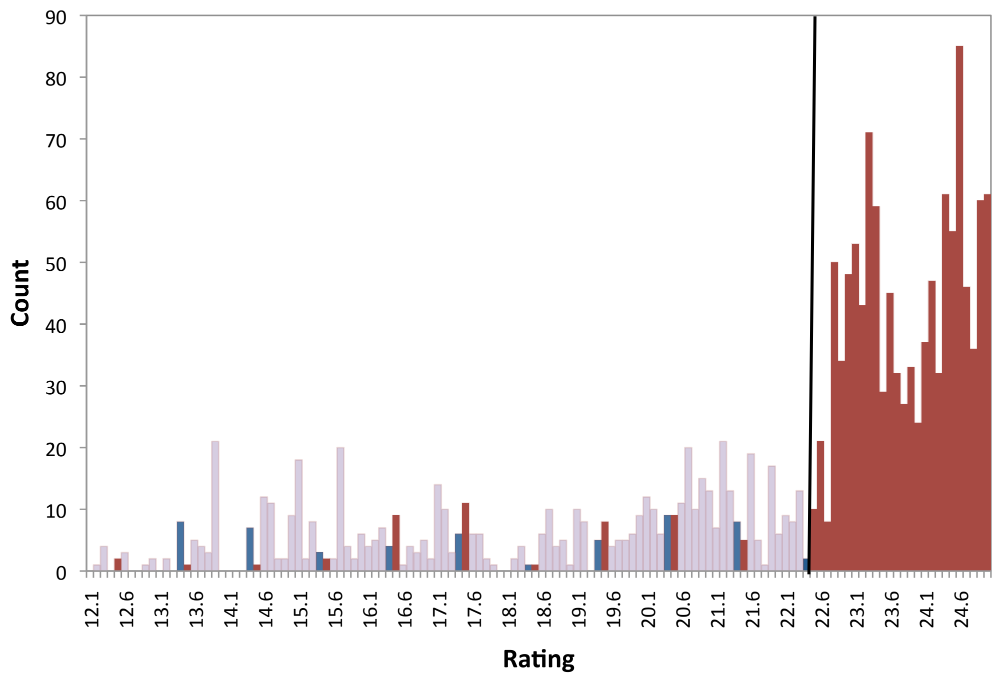
(f) 1983 (Top Notch = 19.0)



Note: Data come from the EPA. Distributions are truncated at 22.5 MPG, but percentages on the horizontal axis reflect the entire distribution, including the portion not shown.

actual number of observations. This regression, which has an R^2 of .81, predicts that 57% of the vehicles within 1 mpg of the top notch will be above the notch – this is more than half

Figure 10: Gas Guzzler Rating Distribution 1998-2007



Note: Data come from the EPA.

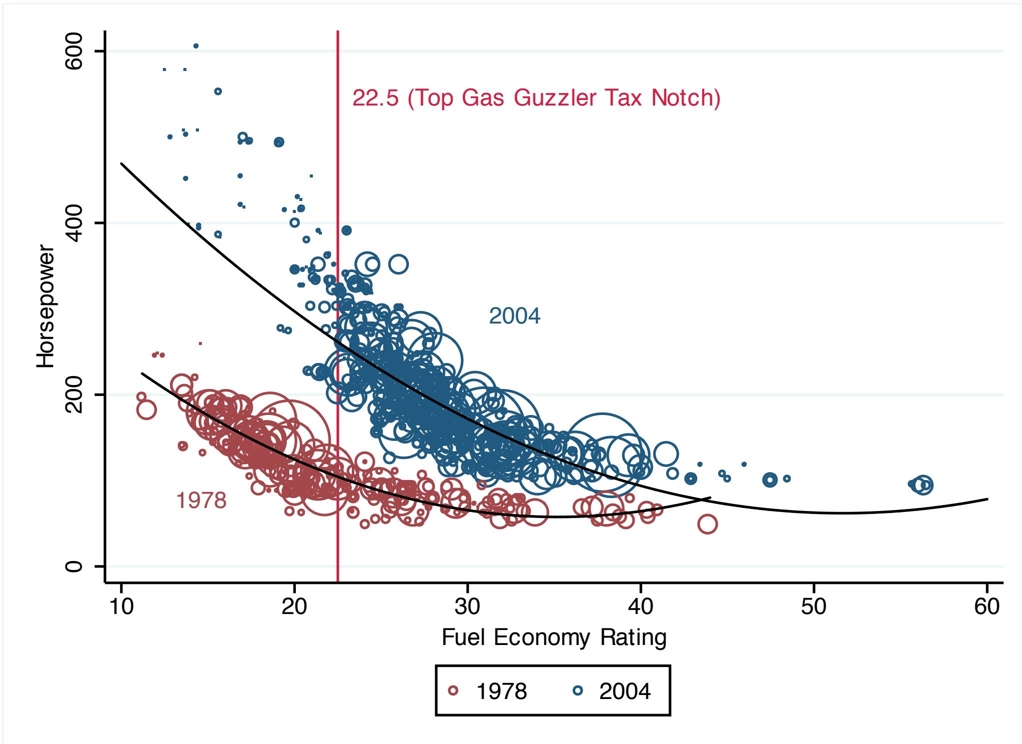
because of the upward slope in the distribution in this neighborhood. In the actual data, 82% of the observations in this window are above the notch. This implies that over two-thirds of the vehicles within 1 mpg of the top notch moved in response to policy, which is equivalent to 3.4% of the vehicles in the entire car market (of whom only a modest fraction are near the notch).

There is nothing special about 22.5 MPG in terms of technology that could explain the apparent discontinuity in the distribution. To be convinced of that, consider figure 11, which is a scatter plot, separately for 1978 and 2004, of models by horsepower and fuel economy rating, where the size of the circles indicate the sales volume of each observation.[20] In each year, there is a negative relationship between the two attributes: other things equal, fuel economy suffers as a vehicle's power increases. The outward shift between 1978 and 2004 indicates technological progress in the intervening years — fuel economy for a given horsepower has increased markedly.[21] For our purposes, what is of interest is the bunching of observations just

[20] Fuel economy data in this plot are taken from CAFE data, where the fuel economy rating is very similar, but not identical, to the Gas Guzzler Tax rating. CAFE and Gas Guzzler Tax ratings differ by around .1 or .2 MPG, so the CAFE rating can be used to show a broad picture but not precise decimal analysis.

[21] See Knittel (2009) for an analysis of this technological shift.

Figure 11: Horsepower and Fuel Economy in 1978 and 2004: Passenger Cars



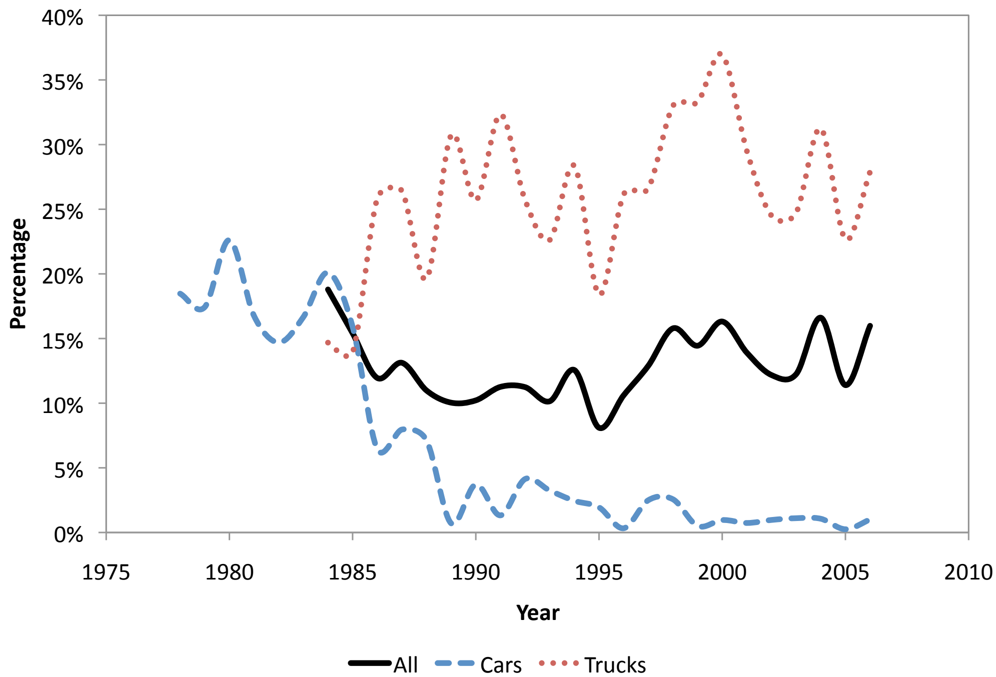
Source: Authors' calculations of NHTSA data.

to the right at the Gas Guzzler Tax threshold of 22.5 MPG, shown by the vertical line. No such bunching is observed in the 1978 data, before the enactment of the Gas Guzzler Tax.

Figure 12, adapted from figure 8 in Sallee (2010), plots the market share of vehicles near the 22.5 MPG threshold since 1978, separately for taxable cars, tax-exempt trucks, and all vehicles. The car series has a precipitous drop-off in that fuel economy class just as the Gas Guzzler Tax began to affect cars with MPG as high as 20, in 1985. In contrast, the market share of trucks in that MPG class jumped sharply at the same time. Sallee (2010) interprets this as evidence that automakers responded to the tax by designing vehicles in that part of the fuel economy distribution to qualify as light-trucks, thereby avoiding the tax and relaxing CAFE constraints.

One reason for the especially large bunching at the top tax threshold is that the tax saving at that notch is significantly larger than for most other notches. Moving over the top notch saves $1,000 (since 1991), whereas other notches are as low as $300. We propose, though, that it may also be that car manufacturers placed a value on avoiding the stigma of a vehicle being

Figure 12: Market Share of Vehicles Rated Between 20 and 22.5 MPG: Passenger Cars, Light Trucks and All Vehicles



Source: Authors' calculations of NHTSA data.

officially labeled a gas guzzler, regardless of the tax liability that came with that designation.

4.5 Evidence from Canada's Feebate

The Canadian feebate provides an opportunity to study automaker response to the introduction of unexpected notches, which requires reaction on a short time horizon. The program was announced in March 2007, at which point all model year 2007 vehicles had been designed and rated and were in the middle of their production cycle. The program appears to have been a surprise to automakers and, even if they had anticipated some policy, it is unlikely that they knew the details sufficiently well to alter fuel economy strategically for 2007 models.[22]

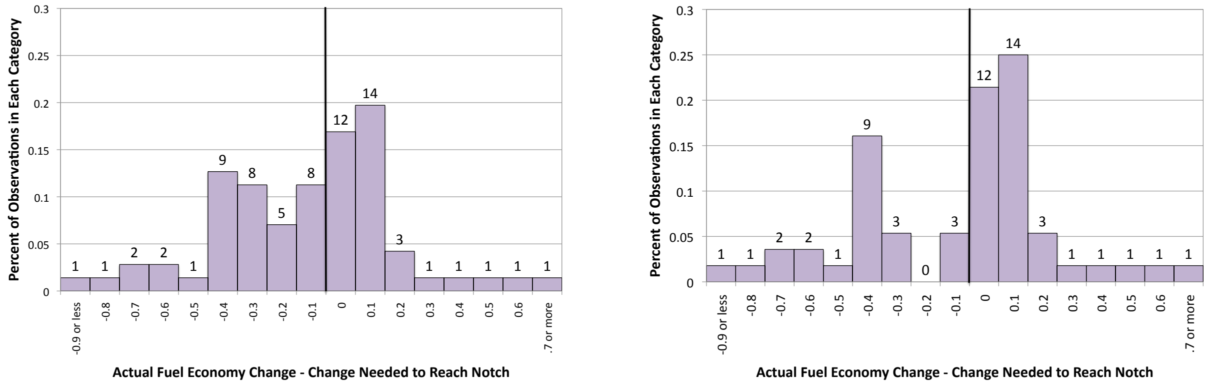
Between March 2007 and the fall of 2007, when most 2008 vehicles began production and received fuel economy ratings, automakers had an opportunity to re-engineer their 2008 model

[22] In 2005, a similar program was discussed but not enacted. Prior to the 2007 budget announcement, there was anticipation that an energy-efficiency subsidy, perhaps including cars, would be included (National Post 2007). But there is no indication automakers were involved in crafting the bill, and they claimed to have been "blindsided" by the policy Keenan (2007b).

Figure 13: Distribution of 2007-8 Fuel Economy Changes Minus the 2007 Distance to a Notch

(a) Vehicles within .3 L/100km in 2007

(b) Vehicles within .3 L/100km in 2007 that Changed Fuel Economy



Note: Total sample sizes are (a) 71 and (b) 56. Raw count shown above each bar.

year vehicles in response to the policy. We expect to see the 2008 version of many models move to the tax-preferred or rebate-preferred side of a notch relative to their 2007 position. To test for this, we matched the official Canadian fuel economy records for 2007 to the corresponding 2008 observation. We declared a vehicle matched if its model name, body type, fuel type, cylinders, displacement and transmission (automatic versus manual) matched across years. We matched 812 models out of 1,040 vehicles from 2007. Many non-matches were due to changes in engine displacement, so our analysis will be biased against finding a response if automakers changed displacement in response to the policy.

In our matched sample of 812 vehicles, 146 models had a non-zero feebate in 2007 – 17 received a rebate and 129 were taxed. Between 2007 and 2008, 766 saw no change in tax or subsidy status. Of the 46 vehicles that changed status, only 5 moved to a less favorable notch, while 41 moved to a more favorable notch. Although this pattern is suggestive of strategic response, secular trends in fuel economy ratings could in principle produce similar changes.

Figures 13a and 13b provide graphical analysis. These figures are histograms of the distribution of the *change* in fuel economy between 2007 and 2008, *relative to* the distance to the nearest notch in 2007. The sample is restricted to models that in 2007 were close to a notch,

Table 5: Regression Test of Response to Canadian Feebate Notches

	Vehicle Improved .1 L/100km or More	Vehicle Improved .2 L/100km or More	Vehicle Improved .3 L/100km or More
Vehicle within .1 L/100km of Notch	0.0372		
	(0.0896)		
Vehicle within .2 L/100km of Notch		0.146**	
		(0.0586)	
Vehicle within .3 L/100km of Notch			0.147***
			(0.0438)
Constant	0.296***	0.194***	0.135***
	(0.0164)	(0.0146)	(0.0130)
Observations	807	807	807
R^2	0.0002	0.0076	0.0138

Note: Standard errors in parentheses. The dependent variables are dummy variables indicating whether or not a vehicle improved its fuel economy by .X L/100km between 2007 and 2008. The right hand side variables are dummy variables indicating whether or not a vehicle was within .X L/100km in 2007.

where "close" is defined as being within .3 L/100k of a more favorable bracket. A value of 0 means that the 2007-8 fuel economy improvement was just enough to make it to the next lower tax or higher rebate in 2008; a -0.1 value means that the fuel economy change fell just short of improving tax treatment. If fuel economy changes did not respond to the feebate notches, we would not expect the density to be particularly high or low near zero, nor would we expect the distribution to be asymmetric. Figure 13a suggests that the opposite is true: there are distinctly more vehicle models just over the next notch than just short of it. Figure 13b replicates this but drops vehicles whose fuel economy stayed exactly the same. Here the asymmetry is much more striking; of those vehicles close to a notch in 2007 that were re-engineered in some way, almost none ended up just short of the notch, while while many ended up just on the tax or rebate-favorable side. Of those whose fuel economy changed, 26 models ended up within .2 L/100km over the notch, while only 3 ended up within .2 L/100km under the notch.

Regression results in table 5 tell a similar story. Each column shows a linear probability model where the dependent variable is whether the change in fuel economy between 2007 and 2008 exceeded a certain threshold, and the lone independent variable (other than a constant) is a dummy variable indicating whether in 2007 the vehicle was within the same value of the next favorable notch. If fuel economy changes are unrelated to notch incentives, a vehicle's proximity to a 2007 fuel economy notch should not affect the probability it improves by a specific amount. The results suggest otherwise. For the second and third column, when the thresholds are .2 and .3 L/100km respectively, being within a given distance to the next notch greatly increases the probability that at least that amount of fuel economy improvement occurs. The third column corresponds to the case shown in figure 13, but effectively compares the response of vehicles within .3 L/100km of a notch in 2007 to all other vehicles, and confirms that the distribution of fuel economy improvement is not typical of vehicles not similarly situated near a notch.

5 Welfare Analysis

Having established that automakers manipulate fuel economy in response to notches, we now develop a framework to assess the welfare implications of this behavior. We remain agnostic as to the potential administrative simplicity or salience benefits of notches, and focus here on documenting the costs that must be outweighed by such benefits in order to justify the use of notched policies. Our framework is built on several assumptions. First, we assume that the policy is motivated by a desire to correct an externality associated with fuel economy and that policy-makers created a notched schedule rather than a smooth one out of administrative simplicity, not because they were targeting the most responsive automakers in the spirit of Blinder and Rosen (1985).[23] We suspect that this is the origin of most policy notches — they are coarse approximations of a smooth schedule rather than strategic decisions.

Second, we assume that automobile design occurs in two stages. In the first stage, which begins many months before production, automakers make global decisions about engine size, body style and vehicle features that have large impacts on fuel economy. At this stage, automakers know a vehicle's fuel economy approximately, but they are uncertain of the exact value and therefore do not know their location vis-à-vis notches. In the second stage, which may be only a few months before production, automakers observe their exact fuel economy by

[23]In reality, fuel economy has social cost implications related to both miles driven, from congestion and accident externalities, as well as gasoline consumed, from air pollution, climate change and energy security externalities. Given our interest in creating a general framework for the welfare analysis of notches, we abstract from these specific concerns and simply assume there is a negative externality associated with lower fuel economy.

testing prototype vehicles and learn their proximity to any notch. During this second stage, automakers decide whether or not to tweak vehicles in response to the notch locations, and we assume this tweaking is limited to less than one full MPG. Moreover, automakers have a limited set of possible modifications, all of which create a discrete fuel economy change. (That these tweaks are discrete is our explanation for why we observe bunching at .6 in addition to .5.) An automaker considering moving over a notch will determine if the cheapest available modification is worth doing. Based on our conversations with engineers, we believe this is a realistic representation of the process.

Third, we adopt a stylized model of both the supply and demand side of the market. On the supply side, we assume that, with respect to individual attributes, automakers behave competitively when determining vehicle design. That is, in the absence of policy, automakers will increase the fuel economy of a given model as long as the cost of doing so is below the consumer willingness to pay, so that if there were no externality, then the provision of fuel economy in the absence of policy would be efficient. We assume that all privately beneficial fuel economy modifications are made in the first stage of vehicle design. This means that policy alone drives second-stage modifications.

On the demand side, we assume that consumers purchase only one vehicle, which avoids the attribute reshuffling described in Kleven and Slemrod (2009), which we believe is peripheral to our analysis here. Consumers value vehicles based on their underlying attributes, one of which is fuel economy. Because of fixed per-variety costs, a finite number of varieties – assumed to be fixed – is produced, and each consumer chooses which variety of car is best, which generates a set of vehicles and a distribution of fuel economy values. Each type of vehicle is optimal for one class of consumers, and given an initial equilibrium, small changes in taxes will induce small changes in vehicle fuel economy, but will not cause consumers to switch vehicles. Thus, the quantity of each vehicle sold does not respond to the fuel economy changes induced by taxes.

Given these assumptions, we can model the introduction of a *smooth* subsidy to increase fuel economy in the market. The social gain from this subsidy SG is approximately equal to $(1/2)e\Delta X$, where e denotes the difference between the marginal social cost and the marginal private cost of MPG which, in the absence of other policies, would equal the externality. The expression ΔX stands for the change in the externality-causing activity, measured in units of vehicle-MPGs.[24] ΔX should be thought of as an industry-wide measure, which averages over the discrete changes made by individual vehicles. Noting that $\Delta X \approx st$, where s captures both demand and supply elasticities, then it is straightforward to show that the value of SG is maximized where $t = e$. This is the first-best outcome with a Pigouvian tax.[25]

If the tax is not equal to the marginal externality ($t \neq e$), the social gain is approximately:

$$SG = \frac{1}{2}(2e - t)\Delta X, \tag{1}$$

which is always less than the case where $t = e$, when (1) reduces to the previous expression. When $t \neq e$, some potential welfare gain is foregone either by failing to induce some socially valuable MPG increase (when $t < e$) or by inducing MPG increases whose marginal social cost exceeds the marginal social benefit (when $t > e$). This is illustrated in Figure 14, where the left panel shows the welfare gain from a too-small subsidy, and the right-hand diagram shows the initial gain, then subsequent loss from setting the subsidy at too high a rate.

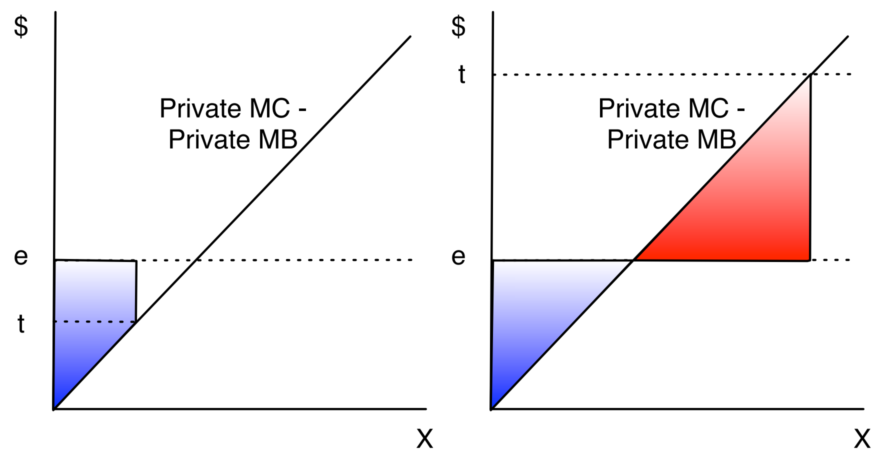
Under a smooth subsidy, all automakers face the same incentive t for each vehicle. Under a notched policy, however, in the second stage some vehicles are close to a notch and others are far. As automakers consider the merits of their least expensive second-stage tweaks, they will face radically different *effective* subsidies per MPG, depending on the proximity of a notch.

To make this concrete, consider the Gas Guzzler Tax, which imposes an average tax of \$800 per MPG, which equates to an \$80 subsidy for every increase of a tenth of an MPG. Suppose that this is the marginal social benefit of increased fuel economy ($e = \$800$), so that a Pigouvian tax would levy an \$80 subsidy per tenth of an MPG. A notched subsidy that jumps by \$800 at a notch every MPG creates an effective per-MPG subsidy that varies depending

[24] E.g., if automakers increase the MPG of one million cars by 0.1 and another 100,000 cars by 0.2, ΔX would be 120,000.

[25] We ignore the social benefit from raising revenue by invoking the argument of Bovenberg and de Mooij (1994) and Bovenberg and de Mooij (1997) that a small environmental tax generates the same labor-market distortion as the labor tax it presumably replaces, and the gross distortion to consumption patterns is small if the tax system is second-best optimal.

Figure 14: Welfare Gains at Varied Subsidy Rates



Note: The blue shaded area in the figure on the left is the welfare gain from a subsidy below the Pigouvian optimum. The figure on the right shows a subsidy in excess of the Pigouvian tax, with the blue area representing a social gain and the red area a social loss. Diagrams represent the entire market of vehicles facing a particular subsidy, and smooth MC - MB curves come from averaging discrete second-stage modification choice across many vehicles.

on the vehicle's decimal after the first stage of design and the size of the jump induced by the second-stage tweak. Table 6 quantifies this variation, where the rows correspond to the initial, "unmanipulated" MPG decimal points, and the columns correspond to ending MPG decimals. Each cell contains the effective per-MPG subsidy from moving from a given starting decimal to a given ending decimal, assuming the jump length is at most 0.9. The magnitude of a jump is measured by the horizontal difference to the right of a diagonal entry marked with an X, including "wrap-arounds" – cases where the ending decimal is below the starting decimal but the integer (not shown) has changed.

The intended subsidy is $800 per-MPG, but table 6 shows that the actual per-MPG subsidy varies from $0 (for tweaks that do not cross a notch) to $8,000 per-MPG (for vehicles starting at a .4 rating with .1 MPG tweaks). The notched system creates several subsidies, some too small and some too big. Although the subsidy is by assumption appropriate on average, the welfare effects deviate from the smooth policy because a set of incorrect taxes that happen to average out to the correct amount do not yield the same welfare gains as a uniformly applied correct tax. It is socially desirable for re-engineering to occur whenever the private cost is

Table 6: Effective Average Subsidy per MPG, Depending on Starting and Ending MPG Decimal

Starting Decimal	Ending Decimal									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	X	0	0	0	0	1600	1333	1143	1000	889
0.1	889	X	0	0	0	2000	1600	1333	1143	1000
0.2	1000	889	X	0	0	2667	2000	1600	1333	1143
0.3	1143	1000	889	X	0	4000	2667	2000	1600	1333
0.4	1333	1143	1000	889	X	8000	4000	2667	2000	1600
0.5	0	0	0	0	0	X	0	0	0	0
0.6	0	0	0	0	0	889	X	0	0	0
0.7	0	0	0	0	0	1000	889	X	0	0
0.8	0	0	0	0	0	1143	1000	889	X	0
0.9	0	0	0	0	0	1333	1143	1000	889	X

Note: Table shows effective subsidy per MPG of a vehicle that starts at the fuel economy decimal in each row value and ends, after manipulation, in the column value. Table assumes a notch value of $800 (the Gas Guzzler Tax average) and includes "wrap-around" values, but assumes all jumps are less than 1 MPG.

less than $800 per-MPG, and to not occur when the cost exceeds $800 MPG. A notch system does not achieve this. For the half of cells with zeros, no reengineering will take place even if the private cost is less than $800 per-MPG. For the half of cells with a positive number, some re-engineering will occur even when its cost exceeds $800; indeed, for vehicles at .4 and jumps of just 0.1 MPG, reengineering will take place as long as the cost is $7999 per MPG or less!

We can use equation 1, which characterizes the welfare gain from an inaccurately sized subsidy, to characterize the social gain of a notched policy by writing the latter as the sum of several of the former:

$$SG = \frac{1}{2}\sum_i (2e - t_i)\Delta X_i = e\sum_i \Delta X_i - \frac{1}{2}\sum_i t_i \Delta X_i = \left(e - \frac{1}{2}\tau\right)\Delta X, \tag{2}$$

where t_i refers to the ith average effective tax, ΔX_i refers to the change in vehicle miles-per-gallon of those subject to it ($\Delta X \equiv \sum \Delta X_i$), and $\tau \equiv \sum t_i \frac{\Delta X_i}{\Delta X}$ is the ΔX_i-weighted average effective tax rate. In terms of Figure 14, this is the summation of a number of these welfare triangles. Note that there is a positive correlation between t_i and ΔX_i (the bigger the tax, the larger the response), which will tend to reduce the value of the net social gain because, for example, the vehicles most likely to have their fuel economy changed are those for which the

Table 7: Hypothetical Distribution for Welfare Calculation

Bins	Average Fuel Economy	Actual Distribution	Counterfactual Distribution
.0	0.05	0.1	0.1
.1	0.15	0.1	0.1
.2	0.25	0.1	0.1
.3	0.35	0.09	0.1
.4	0.45	0.07	0.1
.5	0.55	0.14	0.1
.6	0.65	0.1	0.1
.7	0.75	0.1	0.1
.8	0.85	0.1	0.1
.9	0.95	0.1	0.1

Note: Table shows hypothetical values only.

effective local subsidy per-MPG is ten times higher than the "intended" rate of subsidy. This social gain formula indicates the net social gain of the manipulation that occurs in the second stage of vehicle design.

This formula allows us to use *ex post* data — that is, data on the distribution of vehicles' MPGs after manipulation — and a counterfactual distribution to approximate the welfare consequences of notched local incentives. To build intuition, consider the hypothetical *ex post* distribution shown in table 7. In the absence of manipulation, we expect a uniform decimal distribution, but we observed bunching at .5. Relative to the uniform counterfactual, the data imply that 30% of the vehicles with a starting .4 decimal and 10% of the vehicles with a starting .3 decimal jumped over the notch. We can transform this into a social gain statistic using equation (2):

$$SG = \frac{1}{2}[(2(800) - 8000)(0.1)(0.03N) + (2(800) - 4000)(0.2)(0.01N)] \tag{3}$$

$$= \left(800 - \frac{1}{2}6400\right)(0.005X) = -12N, \tag{4}$$

where τ, the weighted-average effective tax rate, is equal to \$6,400 per-MPG, and N is the total number of vehicles (not vehicle models) in the range of the notch incentives. This says that the net social benefit of second-stage manipulation is negative. Intuitively, this is because the average (weighted by response) effective subsidy of \$6,400 is much larger than the \$800

Table 8: Actual Distribution for Welfare Calculation: Gas Guzzler Tax

Bins	Average FE	Actual Dist.	Counterfactual
.0	0.05	0.1023	0.1
.1	0.15	0.0921	0.1
.2	0.25	0.0989	0.1
.3	0.35	0.0867	0.1
.4	0.45	0.0738	0.1
.5	0.55	0.1260	0.1
.6	0.65	0.1152	0.1
.7	0.75	0.1206	0.1
.8	0.85	0.0840	0.1
.9	0.95	0.1003	0.1

Note: Sample matches figure 5. IRS data; sample size is 1,476.

externality. Most of the vehicle redesign happens for vehicles whose cost exceeds the social benefit of increased fuel economy.

Analyzing the actual distribution of rating decimals for vehicles subject to the Gas Guzzler Tax, shown in table 8, requires the same steps. Because our data show almost exactly 30% of observations fall between .0 and .2 and almost exactly 30% fall between .7 and .9, we assume that all second-stage manipulation took place among vehicles starting at .3 or .4 and moving to .5 or .6.[26] Then, the net social gain is $SG = \left(800 - \frac{1}{2}4766\right)(.0063N) = -\$10.50N$. In the hypothetical example, all of the bunching occurs at .5, so we did not need to make any assumption about which observations started and ended where. In the actual data, there is bunching at .5 and .6, so we must make an assumption about this mapping in order to utilize all the values in the social gain formula. Importantly, our linear approximation of the social gain is *invariant* to the mapping between the *ex ante* distribution and the *ex post* distribution, so long as the counterfactual distribution is correct and second-stage manipulation occurred only in response to the notches and only in sizes smaller than 1 full MPG, as assumed above. This is important because it implies that this welfare statistic can be calculated for any *ex post* data set that reveals bunching, provided a counterfactual is available. Researchers do not need to know which observations started where in order to approximate this social gain statistic.

The local incentives of the notch generates a social loss of $10.50 per car from second-stage

[26]We make a very small adjustment to the *ex post* value at .6 in order to make this assumption hold exactly.

manipulation. The actual *ex post* weighted-average subsidy (τ) of \$4,766 per-MPG is lower than the \$8,000 per-MPG that what would obtain if all manipulation were of length 0.1, but is more than the \$4,000 per-MPG if all the jumps were of length 0.2. As long as this weighted average subsidy is greater than two times the externality, the net impact of the second-stage manipulation will be negative.[27]

Equation 2 calculates the net social gain of all second-stage manipulation, some of which is efficient. Another interesting benchmark is to compare this social gain to the social gain from a smooth subsidy, which would induce all of the efficient modifications and none of the inefficient ones. This requires decomposing our social gain formula into the efficient (the blue triangle in figure 14) and the inefficient (red triangle in figure 14) portions. Elementary geometry provides the decomposition; for a given t:

$$SG = \frac{1}{2}(2e - t)\Delta X = \underbrace{\frac{1}{2}\frac{e^2}{t}\Delta X}_{gain} - \underbrace{\frac{1}{2}\left(1 - \frac{e}{t}\right)(t - e)\Delta X}_{loss}, \tag{5}$$

where the first term is the gain from efficient manipulation and the second term is the loss from inefficient manipulation. Just as with equation 2, each term can be summed across the various t to create a single statistic but, unlike the other measure, the decomposed summations are not invariant to the mapping between starting and ending decimal values.

In our case, the lack of invariance is easily solved through reasonable auxiliary assumptions. If we assume that vehicles only move either .1 or .2 MPG in response to the notch (which seems reasonable given that there is no shortage of observations at .2), then it must be that all of the vehicles that started at .3 ended at .5. This pins down the distribution of ending decimals for those starting at .4, and we can decompose our social gain into the efficient and inefficient portions, which reveals that the efficient gain is only \$.48N, but the inefficient loss is -\$10.98N.

This decomposition puts -\$10.50N into perspective. The net social loss from the second-stage manipulation in response to the notched subsidies is not only negative, but it is also

[27] Also, because there will be a positive correlation between effective subsidy rates and manipulation, τ will exceed the average statutory t, which implies that t, the policy parameter, should be set below the Pigouvian tax that would prevail in a system without notches.

twenty times as large as the benefit from the affected vehicles. Because deadweight loss rises with the square of t and gains are zero when $t = 2e$, it is intuitive that effective subsidies of \$8,000 and \$4,000, which are 10 and 5 times the externality, yield very large inefficiency losses relative to gains. The inefficiency of the notched subsidy compared to the smooth subsidy is even worse if one considers the vehicles that face a zero subsidy under the notch, for which socially efficient improvements are not made. The decomposed social gain from the vehicles starting at the .4 decimal in the data is $\$0.27N$.[28] If all vehicles faced the smooth incentive, then the total benefit of second-stage manipulation would be ten times this amount, or $\$2.70N$. Thus, while the smooth subsidy would generate a \$2.70 social benefit per unit sold for second-stage manipulations, the notched subsidy creates a \$10.50 social loss per unit. If notches exist for administrative or salience reasons, these benefits must be large enough to offset costs of this magnitude.

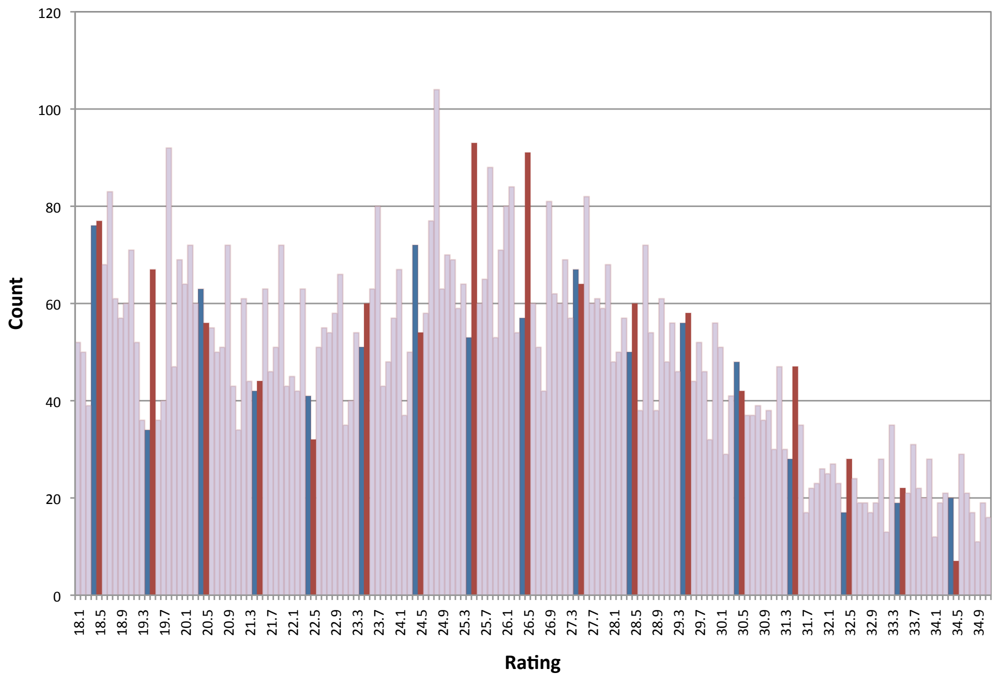
6 Bunching in Fuel Economy Label Ratings

Automakers face fuel economy policy notches not only in the form of tax incentives, but also in the form of fuel economy labels for consumers. Automakers are required by federal law to attach a fuel economy label to all new vehicles, and the contents of this label are strictly prescribed. Ratings must be based on the aforementioned fuel economy test (but are *not* equal to the Gas Guzzler Tax rating), and the values must be reported as integers, where the test results are rounded off. This results in a *presentation notch* at every .5 MPG, the rounding cutoff. If consumers value fuel economy and use the rounded integers when shopping, then automakers have an incentive to manipulate fuel economy ratings around these presentation notches. Whereas only high-performance passenger cars are subject to the Gas Guzzler Tax, all vehicles have labels, which allows us to test for bunching throughout the entire market.

To test for bunching, we generate histograms similar to the ones presented above for the Gas Guzzler Tax. Because the EPA's publicly available data files do not include the unrounded

[28]This follows from using equation 5 on vehicles estimated to have begun at .4 and moved to .5 and .6: $(1/2)800^2/4000(.02).0135 + (1/2)800^2/8000(.01).0127 = .27$.

Figure 15: Highway Label Distribution: 1978 - 1983 and 1998 - 2007



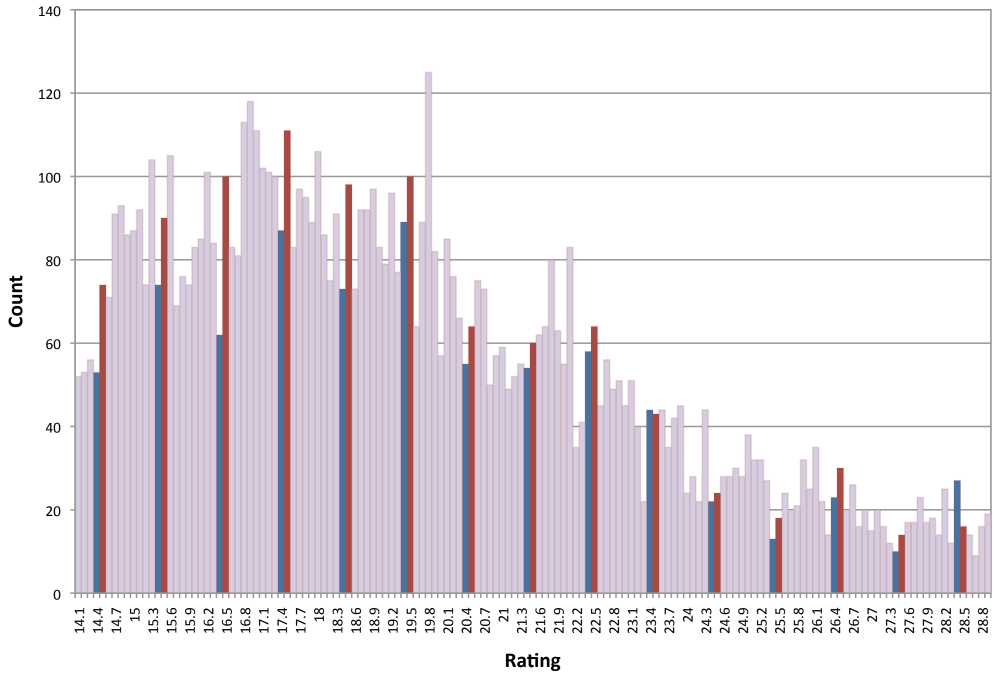
Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

estimates from 1984 to 1997, we are limited to data from years before and after this period. Figures 15 and 16 show, respectively, histograms of highway and city fuel economy label ratings decimals for all vehicles.[29] The cutoff for rounding to the nearest integer is .5, so we expect to see bunching at .5 (where the rating is rounded up) relative to .4 (where the rating is rounded down). Ratings ending in .4 are colored in blue, and those ending in .5 are colored in red.

Figure 15 shows that there are more observations just above notches than just below in highway ratings, but there are several integer values where the reverse is true. Figure 16, though, shows consistent and large bunching above notches in city label ratings. Figures 17 and 18 repeat this exercise for the "Big Three" domestic automakers – Chrysler, Ford and General Motors. Bunching just over presentation notches is amplified in this subsample.

[29]The tabulations in these, and all subsequent figures in this section, exclude a very small number of vehicles with unusual fuel types (e.g., compressed natural gas) that are subject to a different rating procedure, and they drop any observation that appears to be a repeat in the sample such that the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year, with the intention of restricting identical engines that are included in multiple models.

Figure 16: City Label Distribution: 1978 - 1983 and 1998 - 2007



Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

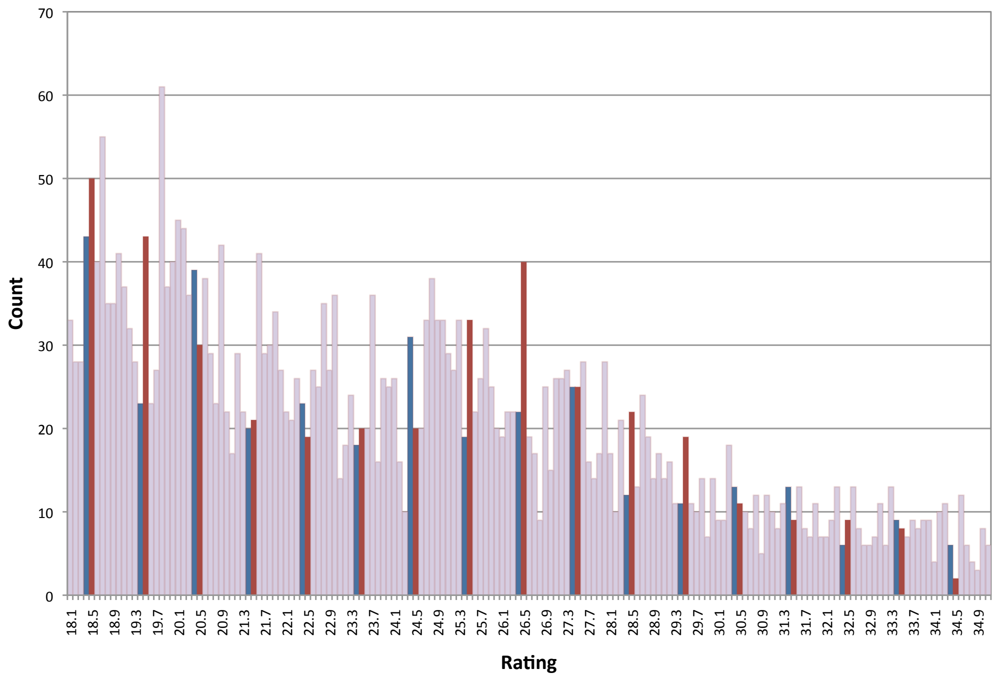
Table 9 shows statistical significance tests for this bunching, repeating the analysis used in table 3 for analysis of the Gas Guzzler Tax. All of the differences are statistically significant at any conventional level, and adjustments for the overall shape of the fuel economy distribution make very little difference in the significance tests.[30]

Figures 19 and 20 show the decimal distributions aggregated across integers, for the full sample and just the Big Three in the late and early year samples separately. These figures show dramatic bunching in the city ratings in both time periods, and there is some evidence of bunching in the highway rating in recent years. Where there is bunching in the overall distribution, it is always greater when the sample is restricted to the Big Three.

There are several reasons why domestic automakers may be more responsive to presentation notches. First, domestic automakers sell a much larger fraction of their vehicles to the U.S. market. Foreign producers may be reluctant to fine-tune vehicles if only a modest fraction will

[30]The appendix includes similar regressions based on statistical tests based on aggregated data – where the fuel economy rating is the unit of observation instead of the vehicle – which confirm the findings here.

Figure 17: Highway Label Distribution, Big Three: 1978 - 1983 and 1998 - 2007

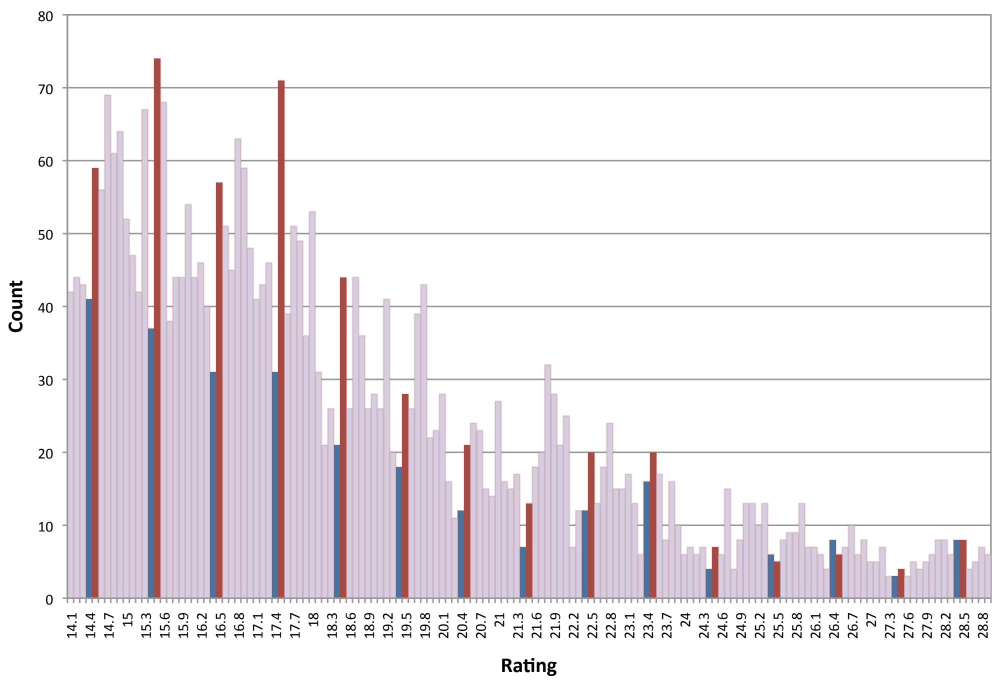


Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

be shipped to the U.S. Second, domestic automakers often face a binding CAFE constraint, which they meet through various strategic adjustments, including fuel economy tweaking. In contrast, European automakers are often out of compliance and pay fines rather than adjust their fleet, and Asian automakers are generally well above the minimum (Anderson and Sallee forthcoming). As a result, the domestic automakers may have developed greater expertise in finely tuning fuel economy to meet CAFE standards. For example, until very recently, only domestic firms made use of a CAFE loophole for flexible-fuel vehicles (Anderson and Sallee forthcoming). Third, relative to American consumers of European cars, American consumers of domestic cars may be more concerned with fuel economy and, relative to the Asian automakers, domestic firms may be more concerned with boosting their fuel economy image.

Finally, figure 21 separates the sample into passenger cars and light trucks, restricting the data to the Big Three. This figure suggests that the city rating shows more bunching for light trucks than cars, whereas the highway rating shows more bunching for cars than trucks. This

Figure 18: City Label Distribution, Big Three: 1978 - 1983 and 1998 - 2007



Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

could be due to the different uses of trucks and cars on average — truck consumers may be more concerned about city fuel economy ratings that reflect the typical usage patterns of larger vehicles, whereas car buyers are more concerned about highway driving and commuting.

7 Inferring Consumer Valuation from Bunching

An important policy question in energy economics is whether or not consumers properly value fuel efficiency in durable goods. If consumers do not properly value fuel economy, then standard results about the superiority of energy taxes to regulatory standards may not hold. This has pivotal implications for energy and climate change policy, as consumer undervaluation may be a justification for CAFE or feebates in addition to a fuel or carbon tax.[31]

[31] Economists generally argue that fuel economy regulation is inferior to a fuel tax because regulation lowers the cost of driving a mile (by raising fuel economy but holding fuel prices constant). If consumers do not properly value fuel economy, however, then they may not respond optimally to changes in fuel prices. If consumer bias is sufficiently large, regulation could be preferred to fuel taxation (Fischer, Harrington and Parry 2007).

Table 9: Statistical Tests of Bunching Above Label Notches

	Observed Number Above Notch (Total Near Notch)	Expected Number Above Notch Under Null Hypothesis	Standard Deviation of Number Above Notch Under Null Hypothesis	p-value
Highway Ratings (Figure 15)				
Binomial Model	791 (1424)	712	18.87	0.00003
Linear Control Function	791 (1424)	709.6	18.64	0.00001
Quintic Control Function	791 (1424)	710.2	18.70	0.00002
City Ratings (Figure 16)				
Binomial Model	742 (1382)	691	18.59	0.00607
Linear Control Function	742 (1382)	690.5	19.20	0.00731
Quintic Control Function	742 (1382)	689.8	19.16	0.00644
Highway Ratings, Big Three (Figure 17)				
Binomial Model	290 (492)	246	11.09	0.00007
Linear Control Function	290 (492)	245.6	11.02	0.00006
Quintic Control Function	290 (492)	245	11.00	0.00004
City Ratings, Big Three (Figure 18)				
Binomial Model	296 (464)	232	10.77	<0.00001
Linear Control Function	296 (464)	232.24	10.17	<0.00001
Quintic Control Function	296 (464)	231.8	10.18	<0.00001

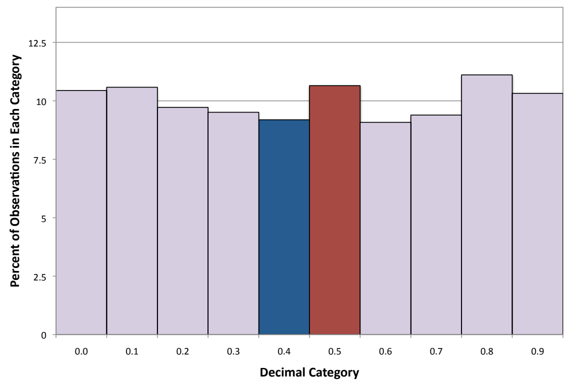
Note: The binomial model assumes that the probabilities that a vehicle has a fuel economy rating ending in .4 and .5 are the same. The linear model modifies this assumption by adjusting for the overall shape of the distribution using a linear fit, excluding observations within .1 MPG of a notch. The quintic model extends this by using a fifth order polynomial to estimate the distribution's shape. Statistics for the linear and quintic models are derived via bootstrapping to account for sampling variation in the estimated polynomial.

There are several reasons to suspect that consumers might undervalue fuel economy. First, experimental evidence suggests that consumers do not understand the nonlinearity in fuel savings from changes in MPG and systematically make mistakes about the relative value of different fuel economy improvements (Larrick and Soll 2008). Second, survey evidence indicates that consumers are unable to articulate the key building blocks for a fuel economy valuation calculation, including mileage, fuel economy ratings and a discount rate (Turrentine and Kurani 2007). Third, the literature on energy-intensive durables has often found evidence of a very large implicit discount rate, which may be a symptom of consumer myopia (Hausman 1979; Dubin and McFadden 1984).[32] Finally, in the car market, researchers have tested whether or not relative prices of efficient and inefficient vehicles adjust as much as the standard models

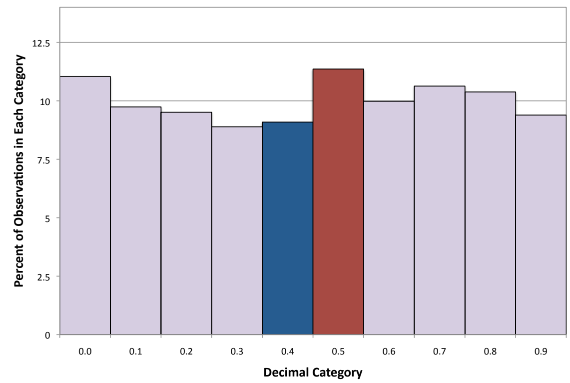
[32] In contrast, Dreyfus and Viscusi (1995) conclude that discount rates are in line with market interest rates.

Figure 19: Label Rating Decimals: 1999-2007

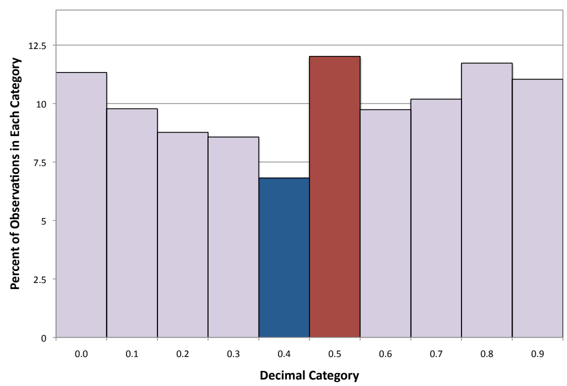
(a) City Ratings



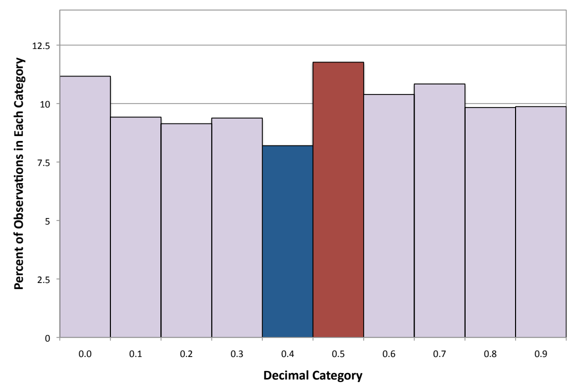
(b) Highway Ratings



(c) City Ratings, Big Three



(d) Highway Ratings, Big Three

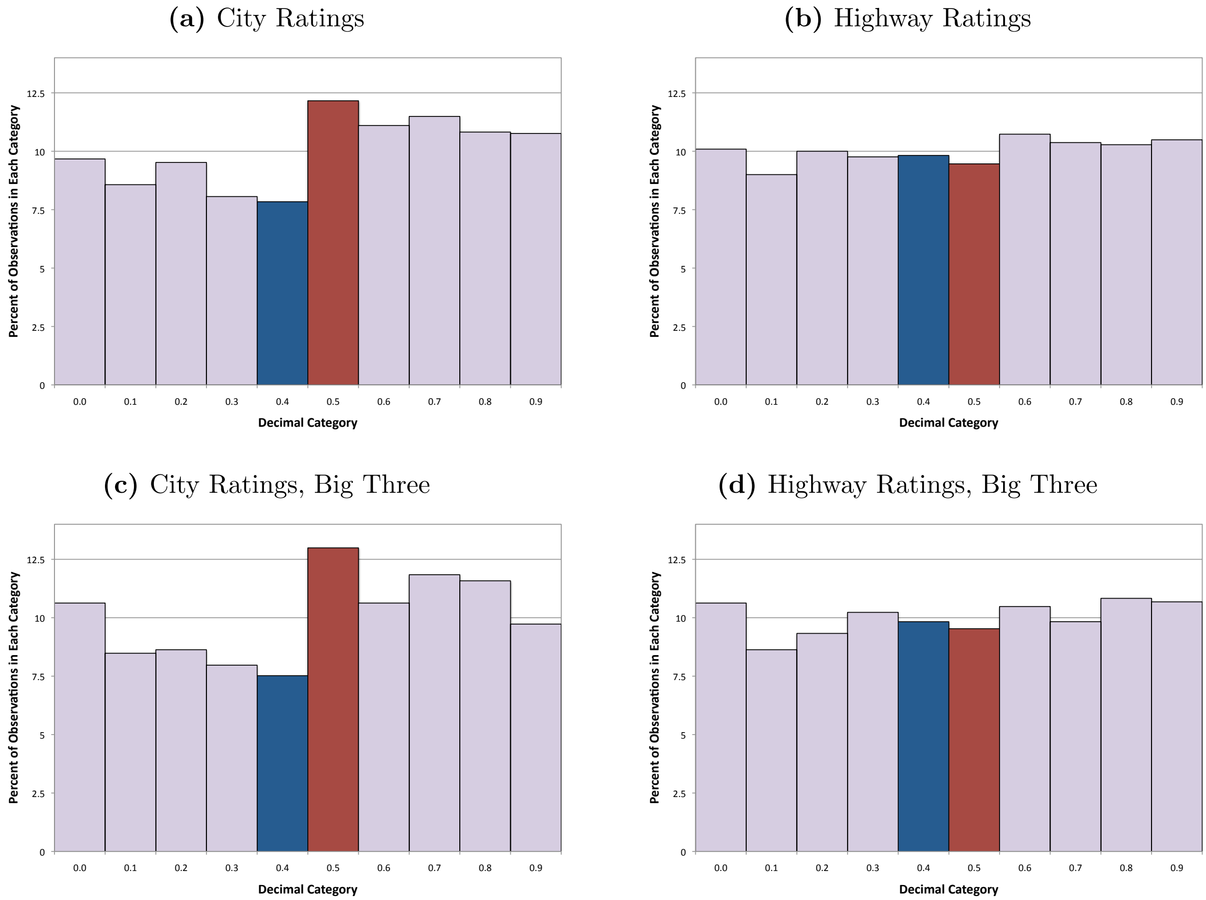


Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

suggest when gasoline prices change. Much of this literature has found that consumers significantly undervalue fuel economy (Kahn 1986; Kilian and Sims 2006; Alcott and Wozny 2009), but some recent work has suggested that consumer valuation may be complete (Busse, Knittel and Zettelmeyer 2009; Sallee, West and Fan 2009). Greene (2010) and Helfand and Wolverton (2009) review this literature and conclude that the evidence is mixed and more research is warranted.

We can combine our analysis of the behavioral response to Gas Guzzler Tax notches with our analysis of the response to label notches to generate an estimate of the consumer valuation of

Figure 20: Label Rating Decimals: 1978-1983



(a) City Ratings

(b) Highway Ratings

(c) City Ratings, Big Three

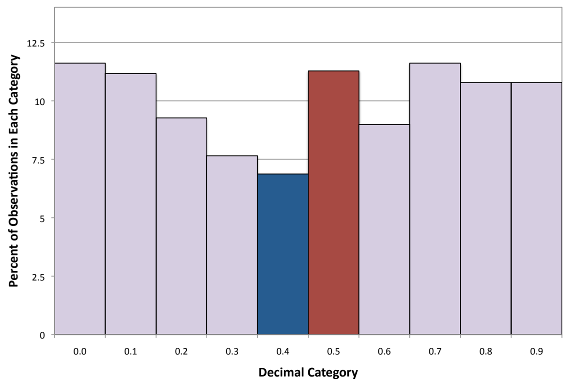
(d) Highway Ratings, Big Three

Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

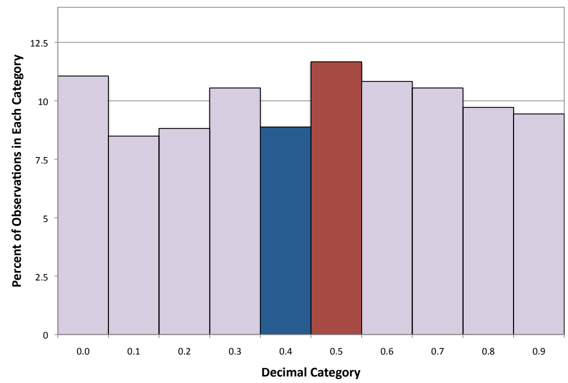
fuel economy. In the case of the Gas Guzzler Tax, we know the dollar value of crossing a notch, because it is set by policy. In contrast, the value of fuel economy to consumers is unknown, so we cannot directly assign a value to fuel economy label notches. But, by comparing the amount of shifting across the two sets of notches, we can *infer* the consumer valuation of an increase in fuel economy. To see this suppose, for example, that all Gas Guzzler Tax notches were worth \$800, and we observed the same amount of bunching around Gas Guzzler Tax notches as we observed around fuel economy label notches. If the cost of manipulation were the same in both cases, then this would imply that a fuel economy notch was also perceived to be worth \$800. If

Figure 21: Label Rating Decimals: 1978-1983 and 1998-2007

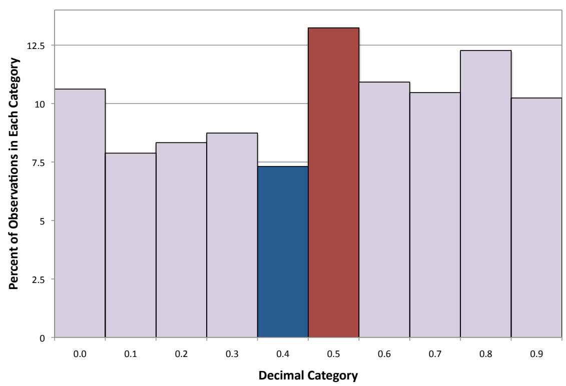
(a) Car City Ratings, Big Three



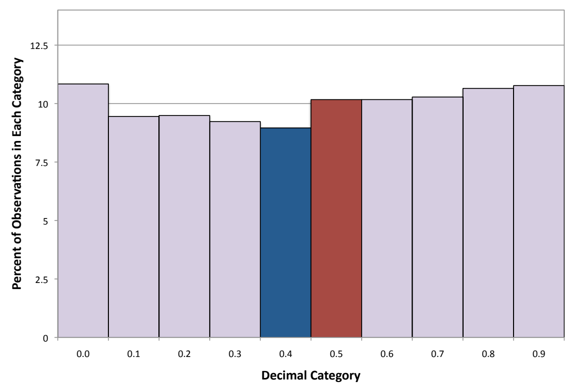
(b) Car Highway Ratings, Big Three



(c) Truck City Ratings, Big Three



(d) Truck Highway Ratings, Big Three



Note: Data come from the EPA. Sample excludes vehicles that run on alternative fuels and observations where the manufacturer, cylinders, displacement, drive type, fuel type, transmission and fuel economy ratings are identical within a model year.

we observed less bunching around the label notches, we would conclude that label notches are worth less than $800, and so on.

The value to the manufacturer and to consumers of crossing a tax notch depends on the ultimate incidence of the tax change. But, as long as the incidence of a tax reduction is the same as the incidence of an incremental attribute improvement, our approach is valid irrespective of incidence. That is, if automakers capture the same proportion of an increased consumer willingness to pay from an exogenous tax reduction as they capture of an exogenous increase in fuel economy, then the notch comparison will shed light on the valuation of fuel economy.

As in section 5, we assume that vehicles are designed in two stages and take the first-stage vehicle design, during which automakers ignore the location of notches, as given. At the conclusion of the first stage, automakers observe that a vehicle's MPG is a distance du from the closest tax-preferred notch and decide whether or not to modify the vehicle in response. We simplify the discussion by assuming that vehicles are located near only one notch.

In the second stage, manufacturers make marginal design decisions about each vehicle i and decide to increase fuel economy if the cost of doing so, denoted $H(du_i, N_i) + \varepsilon_i$ where N is a vector of vehicle characteristics and ε is a random cost component, is less than the benefit of crossing the notch, denoted V. Considering a set of vehicles that have the same du, the fraction of these vehicles that will move over a notch is equal to the probability that $V > H(du_i, N_i) + \varepsilon_i$. Then, a regression of the amount of bunching on the notch value, holding distance and other characteristics constant, reveals the shape of the distribution of the cost function.

We use this insight to estimate the value of fuel economy label notches using the Gas Guzzler Tax regressions from section 4.2 above. To do this, we need to compare the amount of bunching in the Gas Guzzler Tax that occurs within .2 MPG of a notch to the amount of bunching that takes place within a comparable distance of the fuel economy notches. Although all three ratings are based on the same tests, the test results undergo a different transformation for each rating. As a result, a fuel economy improvement of .2 MPG in the Gas Guzzler Tax rating does not necessarily correspond to a change of .2 MPG in the label ratings.

The underlying test statistics for the highway and city tests are transformed via a linear adjustment, as discussed above in section 2, so that $H_L = .78H(u)$ and $C_L = .9C(u)$, where H_L and C_L represent the final label ratings, u represents a measure of underlying fuel economy, and $H(u)$ and $C(u)$ are the raw test scores. The fuel economy rating for the Gas Guzzler Tax is a harmonic average of the two underlying test scores:

$$GG = \frac{1}{\frac{.45}{H(u)} + \frac{.55}{C(u)}}. \tag{6}$$

To determine how a change in u will influence all three measures, we need additional assumptions. We normalize $H(u) = u$ so that $H^{'} = 1$, which is the derivative of H with respect to u. To relate $H^{'}$ to $C^{'}$ we further assume that $C^{'}$ is a constant, so that $\gamma C^{'} = H^{'}$. That is, an engineering change that improves the highway test score by 1 MPG will improve the city test score by γ MPG. Some estimate of γ is required to make our calculations. First, we assume that $\gamma = 1$, so that the fuel efficiency label ratings are affected the same amount by any modification. Second, we assume that $\gamma = .81$, which is the slope coefficient in a cross-vehicle regression of the city rating $C(u)$ on the highway rating $H(u)$ and a constant. This regression yields a prediction equation relating the two measures: $C(\hat{u}) = -6.34 + .81H(u)$ with an $R^2 = .91$.

To compare the amount of bunching across different notch types for a given γ, we calculate the change in u implied by a .2 MPG increase in the Gas Guzzler Tax rating, and then calculate the change in label rating that corresponds to the same u.[33] The results, reported in table 10, are .187 when $\gamma = 1$ and .220 when $\gamma = .81$. These values represent the change in u necessary to move the Gas Guzzler Tax rating by .2 MPG. A change in u of .187 (.220) corresponds to a change in the highway label rating of .146 (.172) due to the in-use shortfall adjustment. The same change in u corresponds to a change in the city label rating of .169 (.160). We then recalculate the amount of bunching in the highway and city label ratings using these values as the width of our bunching window, so that the bunching window for the Gas Guzzler Tax rating, the highway label rating and the city label rating now all represent the same change in u. These bunching estimates are reported for passenger cars in table 10.

To complete the valuation, we calculate the notch value (in 2008 dollars) that would have generated a bunching prediction equal to the observed value using the results from column 6 of table 4, which characterizes the relationship between notch values and bunching. Table 10 shows substantial estimates, ranging between \$402 and \$752. These values can be directly compared to the present discounted value of an increase of one MPG of fuel economy, given

[33] Because the change in u required to achieve a .2 MPG increase in the Gas Guzzler Tax rating depends on the distance between the highway and the city test ratings, there is no single change in u that creates a .2 MPG increase. We calculate the highway and the city rating that would leave a vehicle exactly .2 MPG away from each of the Gas Guzzler Tax notches, assuming that $C(u) = -.634 + .81H(u)$. For each of these starting points, we calculated the increase in u necessary to reach the notch – these values range from .180 to .189 when $\gamma = 1$ and from .214 to .221 when $\gamma = .81$. Then, we tabulate the number of vehicles that are within .2 MPG of each of these notches in the data, and use these tabulations as weights to create a single weighted average.

Table 10: Estimates of Consumer Willingness to Pay for an Additional 1 MPG of City or Highway Fuel Economy: 1998-2007

	$\gamma = 1$	$\gamma = .81$
du_G^*= the change in u required to move Gas Guzzler Tax .2 MPG	.187	.220
du_H^*= the change in Highway Label resulting from change in u of du_G^*	.146	.172
du_C^*= the change in City Label resulting from change in u of du_G^*	.169	.160
B_H= percentage of observations within du_H^* of Highway Label notches on preferred side	.524	.505
B_C= percentage of observations within du_C^* of City Label notches on preferred side	.551	.552
Estimated value of Highway Label notch ((B_H-.450)/.0136)	\$541	\$402
Estimated value of City Label notch ((B_C-.450)/.0136)	\$743	\$752

Note: Estimates are based on data for passenger cars between 1998 and 2007. The estimated coefficient on notch value in specification 6 of table 4 is .0136.

typical driving behavior and average city versus highway mileage. Average mileage is 12,500 miles per year for thirteen years, which we divide between highway and city conditions using the EPA estimate of 55% city and 45% highway. At a gasoline price of \$2.50 and a 5% discount rate, improving a vehicle's city fuel economy from 20.5 (the sample mean) to 21.5 has a value of \$385 dollars, which is similar to our highway estimates, but only half our city estimates. Because the cost saving is a nonlinear function of MPG, this calculation depends on the starting MPG value. An improvement in the highway rating from 28.5 (the sample mean) to 29.5 has a value of \$165, which is much smaller than our estimates.

Overall, the fuel economy valuation implied by the amount of bunching observed between label ratings and the Gas Guzzler Tax ratings implies a substantial consumer valuation of fuel economy. Two important caveats are in order, however. First, the procedure used here does not account for sales volume because we do not have sales data available for the label analysis. Because vehicles subject to the Gas Guzzler Tax tend to be low volume, these estimates may overstate the value of fuel economy. Second, as noted in the discussion of table 4, estimates of the relationship between notch values and the degree of bunching are sensitive to the data source. All of this implies considerable uncertainty around the estimates provided in table 10. Nevertheless, this exercise hopefully serves as an interesting example of how economic parameters can be estimated from studying notches. At a minimum, the fact that automakers bunch in label ratings at all suggests that consumers do value fuel economy.

8 Conclusion

Key aspects of American and Canadian vehicle fuel economy policy are designed with notches, so that many vehicles face no incentive to incrementally improve fuel economy, but others face large and varying incentives for improvement. In this paper we show that the policy notches have real consequences, as there are significantly more vehicles produced (and purchased) just on the policy-beneficial side of the notches than otherwise would be expected. We observe this behavior not only in response to explicit notches in tax and subsidy policies, but also in response to implicit "presentation notches," where government policy dictates what (coarse) information a firm must provide to consumers. We develop a simple framework within which the negative welfare effects of local manipulation can be calculated, a framework which may prove useful in a variety of contexts as it can be utilized with only *ex post* data.

Future fuel economy policies are likely to increase the importance of notches. The state of California recently explore a comprehensive feebate program with notches, and similar legislation has been introduced in the U.S. Senate. The EPA is considering the adoption of a letter grade system with grading notches that would rate vehicles by fuel economy and emissions. Recent CAFE reforms have dramatically tightened its standards in a way that increases the value of moving vehicles over the light-truck classification notch. Policy notches may have administrative or salience benefits, but for notches to be warranted, these benefits must outweigh the substantial inefficiency costs that we document here.

References

Alcott, Hunt and Nathan Wozny, "Gasoline Prices and the Fuel Economy Discount Puzzle," September 2009. Manuscript: Princeton University.

Anderson, Soren T. and James M. Sallee, "Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards," *American Economic Review*, forthcoming.

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

Blundell, Richard, "Work Incentives and 'In-Work' Benefit Reforms: A Review," *Oxford Review of Economic Policy*, Spring 2000, *16* (1), 27–44.

Bovenberg, Lans A. and Ruud A. de Mooij, "Environmental Levies and Distortionary Taxation," *American Economic Review*, September 1994, *84* (4), 1085–1089.

_ **and** _ , "Environmental Tax Reform and Endogenous Growth," *Journal of Public Economics*, January 1997, *63* (2), 207–237.

Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer, "Pain at the Pump: The Differential Effect of Gasoline Prices on New and Used Automobile Markets," September 2009. Manuscript, University of California, Davis.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri, "The Effect of Adjustment Costs and Institutional Constraints on Labor Supply Elasticities: Evidence from Denmark," December 2009. NBER Working Paper no. 15617.

Dreyfus, Mark K. and W. Kip Viscusi, "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Economy," *Journal of Law and Economics*, April 1995, *38* (1), 79–105.

Dubin, Jeffrey A. and Daniel L. McFadden, "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, March 1984, *52* (2), 345–362.

Fischer, Carolyn, Winston Harrington, and Ian W.H. Parry, "Should Automobile Fuel Economy Standards Be Tightened?," February 2007. Resources for the Future Discussion Paper 04-53 (Revised).

Greene, David L., "How Consumers Value Fuel Economy: A Literature Review," Technical Report EPA-420-R-10-008, U.S. Environmental Protection Agency March 2010.

Greimel, Hans, "Nissan Tweaks MPG to Sweeten Payout," *Automotive News*, July 27, 2009.

Harberger, Arnold C., "Tax Lore for Budding Reformers," in Rudiger Dornbusch and Sebastian Edwards, eds., *Reform, Recovery and Growth: Latin America and the Middle East*, University of Chicago Press, Chicago, 1995.

Hausman, Jerry A., "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," *The Bell Journal of Economics*, Spring 1979, *10* (1), 33–54.

Helfand, Gloria and Ann Wolverton, "Evaluating the Consumer Response to Fuel Economy: A Review of the Literature," August 2009. National Center for Environmental Economics Working Paper No. 09-04.

Kahn, James A., "Gasoline Prices and the Used Automobile Market: A Rational Expectations Asset Price Approach," *Quarterly Journal of Economics*, May 1986, *101* (2), 323–340.

Keenan, Greg, "Big Auto's Race to Get Green," *The Globe and Mail*, December 11, 2007.

_ , "'Feebate' Won't Get Guzzlers Off Road, Critics Say; Tax Hasn't Worked in B.C., Ontario," *The Globe and Mail*, March 21, 2007.

Kilian, Lutz and Eric R. Sims, "The Effects of Real Gasoline Prices on Automobile Demand: A Structural Analysis Using Micro Data," April 2006. Manuscript, University of Michigan.

Kleven, Henrik Jacobsen and Joel Slemrod, "A Characteristics Approach to Optimal Taxation and Tax-Driven Product Innovation," September 2009. Working Paper: London School of Economics.

Klier, Thomas and Joshua Linn, "New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard," November 2008. Federal Reserve Bank of Chicago Working Paper 08-13.

Knittel, Christopher R., "Automobiles on Steroids: Product Attribute Trade-Offs and Technological Progress in the Automobile Sector," August 2009. Manuscript, University of California, Davis.

Larrick, Richard P. and Jack B. Soll, "The MPG Illusion," *Science*, June 2008, *320* (5883), 1593–4.

National Post, "Were the Fortune Tellers Right?," *National Post*, March 20, 2007.

National Research Council, *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*, The National Academies Press: Washington, D.C., 2002.

Ramnath, Shanthi, "Taxpayer Reponses to Notches: Evidence from the Saver's Credit," November 2009. Manuscript, University of Michigan.

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

Sallee, James M., "The Taxation of Fuel Economy," October 2010. NBER Working Paper no. 16466.

_ **, Sarah E. West, and Wei Fan**, "Consumer Valuation of Fuel Economy: A Microdata Approach," *Proceedings of the National Tax Association Annual Conference on Taxation*, 2009.

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

Turrentine, Thomas S. and Kenneth S. Kurani, "Car Buyers and Fuel Economy?," *Energy Policy*, February 2007, *35* (2), 1213 – 1223.

Whitefoot, Katie, Meredith Fowlie, and Steven Skerlos, "Evaluating U.S. Reformed Corporate Average Fuel Economy Standards using an Engineering Attribute Selection Model," October 2010. Manuscript: University of Michigan.

Not For Publication

A Additional Regression Tests

An alternative methodology for checking statistical significance of bunching is to collapse the data down to cell counts (e.g., figure 3) and then directly test whether the cell counts are different, taking the counts as the primitive rather than the individual vehicle observations. In the case of the Gas Guzzler Tax, collapsing in this way reduces the number of observations to 100, so this alternative methodology creates a statistical test based on 100 observations, rather than the 1,221 observations that underlie it. For this reason, we think that the methodology employed above is truer to the statistical significance. The cell count regressions therefore provide a harsher test of our hypothesis. For completeness, we provide here a summary of our baseline regression tests on these variables.

Table 11 shows such count regressions for the Gas Guzzler Tax, city fuel economy labels and highway fuel economy labels. Two specifications are included for the Gas Guzzler Tax, where the sales weighted cell counts are the dependent variable. Both include a full set of dummy variables for each decimal rating, with .4 as the omitted category. Only the .5 decimal coefficient is reported for brevity, and it may be interpreted directly as the difference between cell counts for .5 decimal ratings and .4 decimal ratings. The first column shows that the .5 decimals have much larger cell counts in magnitude – the dummy coefficient is larger than the constant, implying that there are more than double the number of sales in .5 decimals than in .4 decimals – but this is not statistically significant. The reason is that there is considerable variance due to the shape of the distribution. Once that is controlled for with the rating polynomial, as in column 2, the coefficient becomes statistically significant (though the magnitude changes little).

Table 11 provides similar cell count regressions for city and highway label distributions. Because sales data are not available in the EPA data, these cell counts are unweighted. In both cases, the sample is restricted to truncate the long tails of the distribution. Just as in the Gas Guzzler Tax regressions, the coefficients are large in magnitude when no polynomial is included,

Table 11: Cell Count Regression Tests for Fuel Economy Ratings Bunching

	Gas Guzzler Tax		City Label		Highway Label	
	Sales	Sales	Count	Count	Count	Count
Decimal = .5	7189	6732***	22.14	22.23***	9.84	9.739*
	(4674)	(2177)	(15.47)	(5.612)	(10.44)	(5.722)
Constant	5768**	3.88E+06	55.57***	-1009	49.88***	55.9
	(2649)	(1.268e+07)	(8.352)	(934.6)	(6.489)	(519.3)
Quintic Control Included?	No	Yes	No	Yes	No	Yes
Decimal Dummies Included?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	100	100	211	211	251	251
R^2	0.064	0.675	0.015	0.902	0.01	0.759

Note: Standard errors in parentheses. The dependent variables are cell counts, and the unit of observation is an individual fuel economy rating (e.g., 18.1). All specifications include a set of dummies for all decimal values, with .4 as the omitted category. Specifications may include a fifth-order polynomial in the underlying fuel economy rating as indicated by the quintic control row.

but they are statistically insignificant. Including the polynomial changes the coefficients little, but makes the estimates statistically significant. Overall, these alternative specifications support our hypothesis that there is fuel economy ratings bunching in response to both tax and label notches.