

Who Benefits When Firms Game Corrective Policies?

Mathias Reynaert

Toulouse School of Economics, University of Toulouse Capitole

James M. Sallee

University of California, Berkeley
and the NBER*

February 10, 2018

Abstract

Firms sometimes comply with externality-correcting policies by gaming the measure that determines policy. This harms consumers by eroding information, but it benefits them when cost savings are passed through into prices. We develop a model that highlights this tension and use it to analyze gaming of automobile carbon emission ratings in the EU. We document startling increases in gaming using novel data. We then analyze the effects of gaming in calibrated simulations. Over a wide range of parameters, we find that pass through substantially outweighs information distortions; on net, consumers benefit from gaming, even when they are fooled by it.

Keywords: gaming, corrective taxation, Goodhart's Law, environmental regulation, carbon emissions, automobiles, fuel economy

JEL: Q5, H2, L5

*A previous version of this paper was circulated with the title "Corrective Policy and Goodhart's Law: The Case of Carbon Emissions from Automobiles." We thank seminar participants at Arizona State, ENSAE-CREST, Fundação Getúlio Vargas, the NTA Annual Conference on Taxation, the NBER Summer Institute, the Stanford Institute for Theoretical Economics (SITE), the Norwegian School of Economics, Resources for the Future, Texas A&M, the Toulouse School of Economics, UC Davis, Yale, CEPR Applied IO conference Jerusalem, ETH Zurich, University of Antwerp, Erasmus University and Heidelberg University for helpful comments. Special thanks to Meredith Fowlie for introducing us to the term Goodhart's Law in this policy context. Sallee thanks support from the USDA National Institute of Food and Agriculture. Reynaert and Sallee thank support from the France-Berkeley Fund.

© 2016 Mathias Reynaert and James Sallee. All rights reserved.

1 Introduction

Sometimes firms comply with a regulation by gaming the measure targeted by policy, rather than by changing their true behavior as intended. This relates to Goodhart’s Law, which posits that “when a measure becomes a target, it ceases to be a good measure.”¹ We study how gaming of this sort can impact consumers who rely on manipulable measures for making choices in a market. Gaming erodes consumer information and induces mistakes. But, when the gaming is done in reaction to a regulatory constraint, firms lower their costs by gaming, which benefits consumers via pass through. The impact of gaming on consumer welfare is thus ambiguous, even when the gaming completely fools consumers, and the net effect depends critically on whether or not gaming is done in response to a policy.

This paper explores the impact of gaming on consumer welfare both theoretically and empirically for the case of automobile fuel consumption ratings. We do three things in this paper. First, we use a novel data set to measure on-road fuel consumption and document gaming of fuel consumption ratings, which escalates dramatically following the introduction of regulations that target this rating. Second, we develop a theoretical model that derives the impact of gaming on buyer welfare in a setting where sellers game energy-efficiency ratings, which buyers use to evaluate products. Third, we conduct welfare analysis using structural estimates of the automobile demand system to quantify the welfare effects identified by our theory in calibrated simulations.

Our empirical analysis considers the introduction of stringent corrective policy in the EU automobile market. Prior to 2007, there were no policies in Europe that hinged directly on fuel consumption ratings, but since then both EU-standards and nation-specific tax schemes have created policy incentives that reward lower laboratory fuel consumption test ratings. To measure gaming, we compare the laboratory ratings, which form the basis of policy, with direct measures of on-road fuel consumption that we construct from a data set that tracks fuel consumption and kilometers traveled for a panel of more than 250,000 drivers over twelve years in the Netherlands. Using these data, we estimate the percentage difference between the laboratory test and on-road performance, which we call the *performance gap*, for each vehicle vintage and model.

We document a sharp rise in the performance gap coincident with policy change. Vehicles produced before 2007 show a small, and relatively stable, performance gap. Vehicles produced after that exhibit a large and rising performance gap, so that 2014 model year vehicles have performance gaps in excess of 50% on average. These results are robust to a number of specification checks and various controls and are similar for all automakers. The rise in the performance gap implies that around 65% of the gains in fuel economy since the introduction of policy, as measured by laboratory tests, are false. Using conventional estimates of lifetime distance traveled and a social cost of carbon of \$40 per ton of carbon, the difference between apparent and actual emissions reductions amounts to \$1.2 billion annually from 2010 to 2014 when extrapolated to all of Europe. We interpret the rise in the performance gap as evidence of gaming in response to policy incentives, in the spirit

¹Goodhart’s original concern was monetary policy (Goodhart 1981). A similar notion, also focused on monetary policy, is captured in the Lucas critique (Lucas 1976).

of Goodhart’s Law, and then turn our attention to our central question—who benefits when firms game corrective policy?—both theoretically and in welfare simulations.

Our theoretical model considers a monopolist who sells a good to a representative consumer. The good has some attribute that is desirable to consumers, but it also creates a negative externality that motivates corrective policy. The attribute closely matches the role of fuel economy ratings in the automobile market.² The attribute is not directly observable, however, so consumer demand and government regulation are based on a measure provided by the seller. The seller can change the measure either by changing the true attribute or by gaming, both of which are costly.

In our model, we allow that some fraction of gaming is undetected by buyers. In the absence of policy, this means that gaming lowers buyer welfare for two reasons. First, gaming causes buyers to mis-optimize (choose the wrong quantity of the good), which leads to a loss in buyer surplus that we call choice distortion.³ Second, gaming causes the seller to raise price because buyers perceive an improvement in the product. This price effect further reduces buyer surplus.

Corrective policy disrupts this logic by flipping the sign of price effects. Regulation raises the cost of production. Gaming allows the seller to lower its costs, and this benefits buyers through lower prices in the same way that a reduction in a tax would. When this price effect dominates choice distortions from faulty information, buyers *benefit* from the seller’s gaming, even when they are fooled by it. We believe that our identification of these two competing effects, which highlights how the impact of gaming hinges on the existence of policy, is unique to the literature on gaming.

Throughout the paper we focus on buyer surplus as a notion of the private surplus of consumers that consider buying the good. This is narrower than consumer surplus which would encompass the externality. We focus on incidence and thus on buyer surplus, but gaming will also impact the level of the externality. The ultimate effect of gaming on the environment will turn on the degree of sophistication of the policymaker, who may increasingly tighten policy to achieve real gains, and of the buyers, who may expand the overall size of the market when they mistakenly perceive lower costs of ownership.

Next, we set out to quantify the price effect and choice distortions in our empirical setting. We demonstrate that the price effects and choice distortions identified by our theory have direct empirical analogs in a discrete choice setting. We then estimate a demand model of the European car market that provides us estimates of consumer preferences and the marginal costs of products. Given these preferences and costs we calibrate the incidence of gaming for a range of alternative assumptions regarding consumer awareness, policy stringency, and the degree of gaming.

We find robust results that align with our theoretical predictions. When there is no corrective

²In the automobile market, fuel economy is a characteristic valued by buyers, and it is directly linked to an externality (carbon emissions). In terms of the model, fuel economy ratings, which are isomorphic to carbon emissions ratings, are the true attribute. But, true on-road fuel economy is observed by neither car buyers nor the regulator. Instead, regulations and consumer-facing fuel economy labels are based on laboratory tests. Firms can improve a test result either by increasing a vehicle’s true fuel economy or by gaming the test.

³The choice distortion that we identify is conceptually identical to the consumer surplus loss due to cognitive frictions or information limitations in several prior papers, including Leggett (2002), Allcott (2013), Sallee (2014) and Dubois, Griffith, and O’Connell (2018).

policy and when consumers are fooled by gaming, we find that lowering perceived fuel costs through gaming leads to modest losses in buyer surplus. A significant majority of these losses come through the price effect. As a result, firm profits rise with gaming, and this comes at the expense of lower buyer surplus.

As suggested by the theory, the welfare effects of gaming change when we introduce a corrective policy. We model a mandated decrease in average fuel consumption ratings that firms comply with via shifting their sales mix towards more efficient models. When firms comply honestly, private consumer surplus falls substantially, as consumers are forced into less desirable products. When we allow firms to relax this regulatory constraint by gaming to meet the standard, we find beneficial price effects for buyers that consistently dominate choice distortions so that the net impact of gaming is to raise buyer surplus. Gaming with or without a policy induces a similar sized choice distortion, but we find that this is an order of magnitude smaller than the price effect for a significant range of parameter choices. These results provide empirical validation for our theoretical prediction: gaming benefits consumers in the presence of stringent policy, even when buyers are fooled. We also show that a corrective policy roughly triples the private benefit to a single firm that games when all others are honest, which implies that policy amplifies competitive pressures that incentivize gaming.

Our paper makes several contributions. First, our finding that carmakers game test ratings in response to fuel economy standards contributes to a large literature that seeks to evaluate welfare implications from fuel economy standards.⁴ Recent articles have pointed out new concerns about the efficiency of fuel economy standards due to interactions with the used car market (Jacobsen and van Benthem 2015), safety (Jacobsen 2013, Bento, Gillingham, and Roth 2017), unintended consequences of attribute basing (Ito and Sallee Forthcoming, Whitefoot and Skerlos 2012), and heterogeneity in vehicle lifetimes (Jacobsen, Knittel, Sallee, and van Benthem 2016). We add to that literature by pointing out the importance of rating accuracy and gaming. Our findings reveal that concerns about gaming prove to be yet another inefficiency of fuel economy standards. Note that our focus is on establishing the incidence of gaming, not on providing a policy evaluation of the EU fuel economy standard. For such an evaluation, see Reynaert (2017), which shows that the standard has induced firms to abate emissions with a mixture of technology adoption and gaming.

The broader second contribution is that we bridge two separate literatures, one on the effects of gaming on market outcomes, and the other on the incidence of externality-correcting policies. In brief, the gaming literature has considered what we call the choice distortion, and the policy evaluation literature has considered what we call price effects. We combine these in our theory model and then determine which one dominates in our empirical analysis.

A theoretical literature on gaming has considered how gaming can impact market outcomes in a variety of models (e.g., Gabaix and Laibson 2006, Heidhues, Köszegi, and Murooka 2017, Frankel

⁴Key studies in this literature include Goldberg (1998), Austin and Dinan (2005), Bento, Goulder, Jacobsen, and von Haefen (2009), Anderson and Sallee (2011), Klier and Linn (2012), Jacobsen (2013), Gillingham (2013), Parry, Evans, and Oates (2014), Whitefoot, Fowlie, and Skerlos (2017), and Durrmeyer and Samano (Forthcoming). See Anderson and Sallee (2016) for a recent review.

and Kartik Forthcoming), but to the best of our knowledge, only Rhodes and Wilson (Forthcoming) and Drugov and Troya-Martinez (2015) consider gaming in the context of regulatory enforcement. In contrast to Rhodes and Wilson (Forthcoming), we consider gaming in response to externality-correcting policy when strong enforcement is lacking. In contrast to Drugov and Troya-Martinez (2015), we include endogenous prices for the good. Empirically, there are a number of papers that show how market measures can be manipulated so as to create market distortions, including test scores (Dee, Dobbie, Jacob, and Rockoff 2016), mortgage markets (Ben-David 2011, Jiang, Nelso, and Vytlačil 2014), or corporate earnings (Roychowdhury 2006). None of these weigh the competing effects of information-based losses and positive price effects on buyer surplus, which is our focus.

On the other hand, many papers have analyzed the incidence of externality-correcting policies, showing, for example, how the costs of complying with environmental regulations are passed through into product prices (Jacobsen 2013), energy prices (Fabra and Reguant 2014), or labor earnings (Walker 2013). We study the flip side of this coin, which is how cost savings from regulatory avoidance can pass through to buyers of a regulated good. We add to that literature by integrating the welfare implications of information erosion into incidence analysis.

Third, we also contribute to a strain of environmental economics that studies compliance and enforcement issues.⁵ Environmental economics often cites ease of enforcement and compliance differences as a reason to prefer some instruments over others (e.g., Goulder and Parry 2008). We point out differences in the incidence of alternative instruments (e.g., fuel economy standards versus a gasoline tax) that arise from gaming, which can break the irrelevance of statutory incidence in determining economic incidence. Our point about statutory incidence is also made in Slemrod (2008) and Kopczuk, Marion, Muehlegger, and Slemrod (2016), though those papers do not discuss externality correction.

Fourth, our analysis is also relevant to the specific issue of emissions testing reliability, which has become a major policy topic in the wake of the Volkswagen emissions testing scandal. Our reduced-form estimates of the degree of gaming are consistent with a growing set of media accounts, with analysis performed by the International Council for Clean Transportation, which has documented a similar gap in a number of European countries (The International Council on Clean Transportation 2014, 2015), and with quasi-experimental evidence on fuel consumption gaming in Japan (Tanaka 2017).⁶ We add further evidence of gaming to that body of work, and also introduce the first analysis of how emissions test gaming impacts the private welfare of car buyers. Our detection of gaming via the comparison of two data sets adds to a related tradition in forensic economics (see Zitzewitz (2012) for a review).

⁵See Gray and Shimshack (2011) and Sigman (2012) for reviews, and Dufo, Greenstone, and Ryan (2013), Oliva (2015) and Sanders and Sandler (2017) for related examples.

⁶The ICCT reports include analysis of the Netherlands based on the same data that we use here. Our results are broadly similar, but our analysis contributes by providing an independent academic analysis and through a number of methodological improvements, including the calculation of standard errors, the introduction of time of driving fixed effects and other controls, a variance decomposition, an Empirical Bayes correction, and extrapolation to the entire European market through matching.

Finally, recent research has documented a failure of energy efficiency programs to realize ex ante expected savings in appliances (Davis, Fuchs, and Gertler 2014) and buildings (Fowlie, Greenstone, and Wolfram Forthcoming, Levinson 2016). We provide related evidence of a lack of efficacy in energy-efficiency policies for automobiles, which have delivered only a fraction of expected savings because of gaming.

The paper is structured as follows. We begin with our reduced form analysis of the performance gap in section 2. In section 3 we describe our theoretical model, which delineates choice distortions and price effects. Section 4 estimates a structural model of the automobile market and uses the estimates to calculate changes in consumer surplus due to choice distortions and price effects from gaming in our calibrated simulations. Section 5 concludes.

2 Automakers game EU fuel consumption ratings

In this section we discuss Goodhart’s Law in the EU automobile market. We document a sharp decrease in the correlation between official and on-road fuel consumption ratings that coincides with those ratings becoming subject to regulation. We start by describing policy changes in the EU automobile market and detailing how carmakers report official fuel consumption ratings. We then describe data from a large sample of drivers in the Netherlands that we use to construct estimates of on-road consumption. Finally, we quantify the gap between official ratings and on-road consumption and relate that pattern to the rollout of policy.

2.1 European fuel economy regulations are relatively new

Prior to 2007, automobiles in Europe were not subject to explicit regulation based on their carbon emissions. Since then, automobiles have become subject to both an EU-wide regulation and nation-specific tax and subsidy schemes that are explicitly tied to carbon emissions. The EU-wide regulation mandates that automakers sell vehicles that have a sales-weighted average carbon emission rate below a certain level. A vehicle’s carbon emission rate is measured as a linear transformation (to convert units) of fuel consumption, which itself is just the inverse of fuel economy. Thus, while the EU policy is described as regulating carbon, it is directly analogous to the US Corporate Average Fuel Economy standards, which are described as regulating fuel economy.

Prior to 2007, there was no legally binding fuel economy regulation in Europe. The standard was announced in 2007, passed into law in its final form in 2009, and phased in over several subsequent years. The first year of enforcement was 2012, with a ramp-up in the standard taking place from 2012 to 2015. Fully phased-in, the regulation is quite aggressive by historical and international standards. Fines for non-compliance are stiff. When the standard was announced, if automakers had left their fleets unchanged, they would have faced an average fine of 1,250 euros per vehicle.

In addition to the EU standard, all member states have separate tax schemes for new vehicles. During the rollout of the EU-wide regulation, many member states adjusted their national policies so that tax schemes were based explicitly on carbon emission rates. France introduced a feebate, which

taxes heavily-polluting cars and subsidizes cleaner models, in 2008. In 2009, Germany switched its annual road taxes so that they depend on carbon emissions rates, rather than engine cylinders. In 2008, Spain introduced registration taxes that depend on emissions ratings. The Netherlands did the same in 2010. Gerlagh, van den Bijgaart, Nijland, and Michielsen (Forthcoming) present a detailed overview of the move towards carbon taxation in EU member states after 2007.

2.2 How are fuel economy ratings measured?

Carbon emissions ratings for all of these policies are based on a laboratory test, called the New European Driving Cycle (NEDC). This test procedure, which is conducted by third-party facilities that are funded by the automakers, measures fuel consumption in liters per 100 kilometers (L/100km). A test vehicle is put onto a chassis dynamometer (a treadmill for cars), and a professional driver “drives” the car through a specified series of speeds and accelerations. Emissions are captured directly from the tailpipe and used to determine gaseous outputs, which are used to determine fuel consumption. Two different test cycles are run to simulate city and highway driving. A coast down test is also performed to measure aerodynamics.

The NEDC is not only the basis of CO₂ regulation, but also consumer-facing labels and emissions limits on local air pollutants, such as NO_x, PM and CO. The test procedure captures local pollutants and measures their quantities to determine vehicle compliance with emissions limits. In terms of consumer information, the NEDC rating is the rating that automakers are required to use in consumer advertising, and it is the rating that appears on mandatory energy efficiency labels for new vehicles. As such, the NEDC was used as a regulatory and market instrument for roughly a decade before the rollout of carbon regulation.

We are interested in how automakers might game this test. Tests are performed in third-party facilities, but these are funded by the automakers and may not have incentive to maintain test integrity. According to media and industry accounts, the European test procedure offers the tester considerable “flexibility” in test procedures.⁷ For example, automakers are not expressly prohibited from submitting test vehicles that have been modified. Firms remove optional equipment, thereby changing the weight of the vehicle to improve performance. They also tape down seams in the vehicle, remove side mirrors and roof racks, and over-inflate tires to improve results of the coast down test. Alternatively, automakers may install technologies that perform particularly well on the test cycle, or they could even calibrate an engine to perform in a particularly efficient way during the highly specific test cycle’s series of speeds and accelerations.⁸ To detect gaming, we compare laboratory test ratings with a direct measure of on-road fuel consumption, which we describe next.

⁷For example, see “Europe’s Auto Makers Keep Test Firms Close” in the March 21, 2016 *Wall Street Journal*: <http://on.wsj.com/1o5h47B>.

⁸These “flexibilities” differ significantly from analogous tests in the US, where the law is much more explicit about the details of the test vehicle and test procedures.

2.3 On-road fuel consumption data

We obtained data from TravelCard NV, a company providing fuel services in the Netherlands. These panel data contain information on 66 million fuel station visits from drivers using a TravelCard NV card between January 2004 and May 2015. Most of the individuals in this sample drive a vehicle provided to them by their employer, who also pays directly for fuel. This implies that we have a selected sample, though the provision of a company car is quite common in the Netherlands due to tax advantages and the high cost of personal vehicle ownership. Nearly half of new vehicles sold in the Netherlands are registered as company cars (Booz & Company 2012).

When visiting a fuel station, Travelcard NV users swipe a smart card to pay for fuel. When a driver swipes her card we observe the drivers' license plate and the date, time and location of the fuel station visit. We also observe the exact amount and the type of fuel purchased and a self-reported odometer reading at the time of fueling. Regarding the latter, drivers are asked to enter their current odometer reading into the fuel pump's keypad during the transaction. Transactions are linked to an account, which is a unique combination of a driver with a particular vehicle, which corresponds to a unique license plate. Unfortunately, if the same individual drives one vehicle for some time and then switches to another vehicle, we have no way of linking those data and must treat them as separate accounts. We refer to this panel variable as a driver, which should be understood as a unique combination of a driver with a particular vehicle.

TravelCard NV provides us with a second dataset that matches each license plate with the vehicle brand, model name, weight, fuel type and the official fuel consumption rating of the vehicle. These characteristics allow us to match the Dutch data with a panel on European car sales and prices from 1998-2011 used in Grigolon, Reynaert, and Verboven (Forthcoming) and Reynaert (2017). We match each vehicle to sales volume, list price and a broader set of characteristics, including length, width and several engine characteristics.

Odometer readings are self-reported and many individual entries are unreliable. In appendix A.1 we detail steps we take to purge unreliable entries. Our final sample includes 24 million transactions. These data include over 2,500 unique types of cars driven by 266,000 different drivers. A car type is defined as a unique combination of brand (Volkswagen), model name (Golf), fuel type (Diesel) and official fuel consumption. We define the release year for a car type as the first time we observe a unique combination of these variables in the data between 1998 and 2014.

Table A.1, in the appendix, gives summary statistics for the raw data. The average vehicle in our sample has a fuel consumption of 6.65 L/100km and a weight of 1,354kg. Somewhat less than half of the vehicles (46%) have diesel engines. We observe an average of 107 drivers per car with a maximum of 3,228 drivers. For each of these drivers we observe an average of 134 visits to the pump with a maximum of 1,135. Driver mean total consumption is 6,015 liters of fuel purchased corresponding to 111,726 km travelled. Finally, the average fuel station visit involves 45 liters of fuel purchased, corresponding to an odometer increase of 671 km with a standard deviation of 192 km.

Our sample is drawn from company cars, but our data provide estimates that characterize a

majority of the market. In appendix Table A.2, we show that vehicles in the Travelcard data are cheaper, lighter, more fuel efficient and more likely to be a diesel than the average new vehicle in the Netherlands. One explanation of the difference is that our data contain almost no luxury vehicles or sports cars. Also, company cars are driven more than the average household car, which likely explains the higher share of diesels in our sample.

Nevertheless, the models in our final sample are the models that account for 76% of all sales in the Netherlands. Moreover, the Dutch car market is very similar to the entire EU market; it features a mix of French, German and foreign brands that is very similar to the EU average. Thus, our data provide estimates of on-road performance that cover the broader market fairly well. However, we wish to stress that we have no way of directly assessing whether drivers of company cars drive their vehicles differently, and thereby exhibit a different performance gap, than the average Dutch driver. Fortunately, we are focused on *changes* in the performance gap over time, so as long as differences between company car drivers and other drivers are stable over time, this will not be a major concern for our conclusions.

2.4 On-road fuel consumption and the performance gap

We construct a measure of on-road fuel consumption r_{nij} for each pump visit n of driver i in car j as the ratio of the liters purchased and the change in reported odometer between the visit and the previous visit:

$$r_{nij} = \frac{\text{liters}_n}{\text{odometer}_n - \text{odometer}_{n-1}} * 100. \quad (1)$$

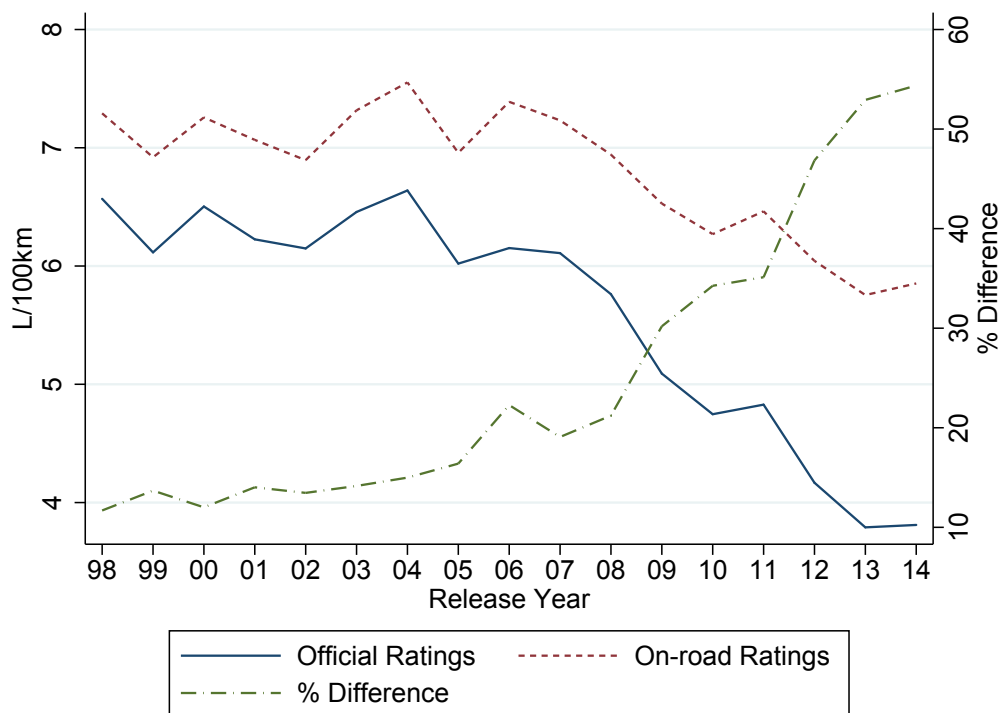
This measure of on-road fuel consumption, in units of L/100km, will vary between pump visits of a driver for three reasons. First, variable driving conditions such as outside temperature, route choice, driving style and congestion will differ across observations. Second, the driver may over or understate the odometer reading. We are not aware of any incentive for the drivers to deliberately misreport distance traveled, but from the data it is obvious that there are many mistakes. Third, there might be variability due to stockpiling effects. If the consumer does not always fill the tank of the vehicle completely there will be variation in r_{nij} . If a driver visits the fuel station with an empty tank and fills half of it we will observe a very low fuel consumption for visit n and a higher fuel consumption for the next visit if she refills the tank completely.

Next, we construct the percentage gap between on-road and tested fuel consumption as:

$$d_{nij} = \frac{r_{nij} - li_j}{li_j}, \quad (2)$$

in which the official rating li_j is constant for each car j and the on-road rating varies across observations. We call this the performance gap. We are interested in estimating the mean and variance of d_j , defined as the average d_{nij} across n and i for a given vehicle type j . In particular, we are interested in the mean and variance of d_j across vehicle types from the same vintage (release year). We now discuss this performance gap by presenting its evolution over time, by estimating a

Figure 1: On-road and official fuel consumption per release year



fixed effects model and by analyzing heterogeneity in the gaps over time.

2.5 The mean performance gap rises over time

Figure 1 plots the mean official rating l_{ij} , the mean on-road rating r_{nij} and the percentage-gap d_{nij} for each release year. Between 1998 and 2006 we see that both official fuel consumption and on-road consumption vary between 6.1 and 7.5 L/100km. The percentage gap fluctuates between 12% and 20% and shows evidence of a slight upward trend after 2004. From 2006 onward we see a spectacular drop in official consumption from 6.2 to less than 4 L/100km. This translates to a rise from the already high value of 38 mpg in 2004 to a truly remarkable 67 mpg in 2014. Official fuel consumption decreases by almost 50% over the sample period.

The on-road ratings follow a trend similar to the official ratings up until 2008. After 2008 the on-road fuel consumption decreases much more slowly than the official rating, going from 6.5 to 5.9 L/100km. As a result, the percentage gap between the official and on-road ratings increases dramatically: from around 12% at the beginning of the sample to almost 55% by 2014. This divergence is remarkable in magnitude, and it coincides exactly with the new EU-wide fuel consumption regulation, which was announced in 2007, finalized in 2009, and phased-in over the remainder of the sample, as well as the phase-in of national policies described above.

The divergence in test ratings and on-road performance is not isolated to a particular automaker. Figure 2 plots the estimated mean gap d_{nij} for three sets of model years (early, middle and late)

Figure 2: Gap between on-road and official fuel consumption per firm

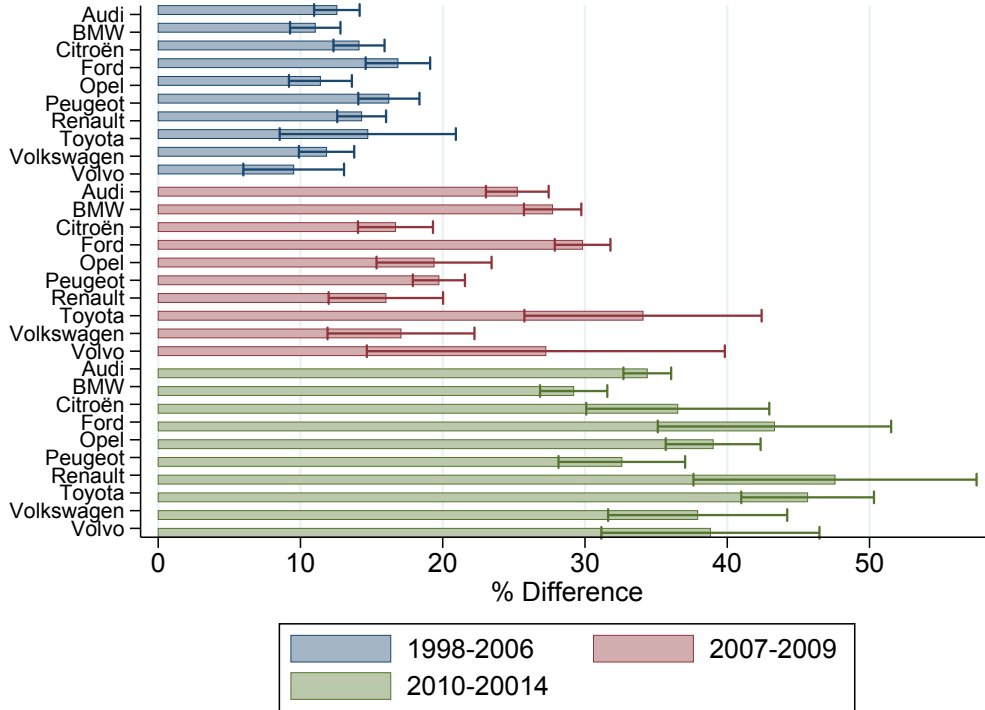


Figure shows estimated coefficients and standard errors from regressing the performance gap on three sets of model release years (early, middle and late) per automaker, the data is restricted to the ten brands with most observations.

separately for each automaker. All automakers show a substantial increase in the gap over time, and all show an economically important gap in the later years. The performance gap is a global phenomenon: the three largest gaps are for a European firm (Renault), a Japanese firm (Toyota), and an American firm (Ford).

2.6 The trend in the performance gap is robust to controls

To account for various potential confounders, we next estimate the performance gap through a series of regressions on our microdata that take the following form:

$$d_{nij} = RELEASE_j + FUEL_j + MODEL_j + YEAR_n + MONTH_n + \epsilon_{nij},$$

where $RELEASE_j$ are release year (i.e., vintage) dummies; $FUEL_j$ is a dummy for diesel engines; $MODEL_j$ are model dummy variables; and $YEAR_n$ and $MONTH_n$ are time fixed effects for year and month, specific to the observed station visit.

We interpret the release year dummy coefficients as average measures of the performance gap for each vintage. These are our coefficients of interest. The omitted category is vehicles that are present in the first year of our data, which implies that they were released in 1998 or before.

If some vehicle types have a bigger gap than others, then our time trend could reflect compositional changes. Therefore, we add a fuel type dummy for diesel and vehicle model name dummies to control for compositional changes in the car market over time. Controlling for model dummies, which is more aggregate than our vehicle type j (e.g., a Toyota Camry is a model, whereas a 2010 Toyota Camry with 2.0L engine is a vehicle type), isolates variation between different engine releases over time (across vintages) of the same model.

Year and month of driving dummies control for driving conditions that change over time and over months in a year. These time controls account for changes in driving behavior, sample selection and other factors such as fuel prices that vary over time by including time of driving fixed effects. For example, we observe both a 2008 and 2009 vehicle type being driving in 2010, and so we can estimate vintage effects (2008 versus 2009) while controlling for time of driving (2010).

Also, we observe later year vintages disproportionately in later years, so changes over time in congestion, driving patterns, or fuel prices could potentially bias our results if we omit time of driving controls.

Figure 3 plots the coefficients and their 95% confidence intervals (with clustering on vehicle type j) from two regressions with individual refueling transactions as the unit of observation. In the first regression (red markers) we estimate the percentage-gap d_{nij} as the coefficients on release year dummies. The second regression (blue markers) adds the controls for time of driving and model fixed effects. This shows that the introduction of controls has little impact on our results, which suggests that compositional changes or changes in driving conditions do not explain the time trend. The same coefficients are presented in columns (1) and (2) of Table 1.⁹

Our interpretation of this time trend in the performance gap is that the accuracy of official fuel consumption ratings declines over time, due to the introduction of European fuel economy regulations. In Table 1 we show evidence that weighs against several alternative explanations. First, we consider the role of a rebound effect. If more recent vintages are more fuel efficient, consumers may respond by driving less carefully or using temperature controls or other equipment more often. In this case, reduced average fuel consumption rates will lead to an increasing gap. One might expect the same mechanism to create a significant difference in the gap between gasoline and diesel powered vehicles, as diesel vehicles are about 30% more energy efficient. If the rebound effect were a major driver of the performance gap, we would expect to see a bigger gap among diesels that grows more quickly. We estimate the release year fixed effects separately for gasoline and diesel engines in columns (3) and (4) of Table 1, and find no such pattern. There is initially no difference in the gap between diesels and gasoline vehicles, in some years diesels have a larger performance gap, while in other years the reverse is true.

Second, one might be concerned that fuel consumption differs with the age of the car. One might expect older cars to become less fuel efficient, leading to an overestimate of the earlier release year dummies. There could also be sample selection, however, such that cars with good on-road fuel

⁹Note that the R^2 for the model without controls is .28, which suggests that release years explain an impressive fraction of the variation, given that the unit of observation here is an individual fuel transaction.

Table 1: Release year fixed effects from regression of performance gap on controls

	All (1)	Controls (2)	Gasoline (3)	Diesel (4)	1st Year (5)	Long (6)
Constant	11.69 (0.66)	13.59 (0.60)	11.44 (0.67)	11.94 (1.13)	10.28 (1.39)	7.80 (0.61)
1999	2.00 (1.45)	0.50 (0.88)	0.96 (1.10)	2.33 (2.10)		2.45 (1.29)
2000	0.32 (0.88)	0.73 (0.73)	0.13 (1.21)	0.40 (1.33)		0.38 (0.88)
2001	2.33 (0.98)	-0.20 (0.96)	2.08 (1.16)	2.40 (1.50)		2.70 (1.02)
2002	1.77 (1.56)	1.28 (1.07)	2.40 (1.50)	1.33 (2.32)		1.48 (1.48)
2003	2.46 (1.32)	1.35 (0.80)	3.79 (1.60)	1.16 (2.03)		1.49 (1.46)
2004	3.30 (1.41)	2.60 (0.78)	3.93 (2.69)	2.80 (1.56)		1.24 (1.33)
2005	4.71 (1.43)	4.29 (0.97)	8.55 (1.84)	2.54 (1.99)	2.50 (2.39)	3.60 (1.25)
2006	10.59 (5.09)	3.93 (1.08)	14.71 (5.96)	3.40 (1.67)	5.80 (1.73)	7.57 (4.52)
2007	7.38 (1.39)	8.55 (1.04)	10.75 (1.50)	5.02 (1.92)	2.79 (2.09)	6.02 (1.21)
2008	9.53 (1.37)	9.00 (1.00)	10.78 (1.09)	8.53 (2.33)	12.60 (2.47)	6.21 (1.20)
2009	18.50 (2.24)	13.92 (1.42)	19.14 (2.94)	17.61 (3.18)	15.66 (2.21)	13.65 (1.89)
2010	22.57 (2.10)	18.13 (1.43)	16.85 (1.42)	25.36 (2.80)	19.85 (2.09)	18.17 (2.11)
2011	23.43 (1.57)	18.54 (1.11)	23.10 (3.64)	23.54 (1.42)	23.28 (2.10)	18.29 (1.24)
2012	35.16 (2.95)	28.81 (2.07)	27.51 (2.17)	40.81 (2.52)	36.96 (4.38)	29.55 (2.94)
2013	41.24 (1.89)	36.06 (1.64)	35.26 (1.97)	43.38 (1.90)	43.16 (2.48)	38.58 (1.80)
2014	42.68 (2.47)	40.92 (3.49)	39.98 (1.69)	48.99 (2.46)	41.50 (2.52)	46.25 (3.84)
Year/Month F.E.		Yes				
Fueltype		Yes				
Model F.E.		Yes				
#Obs. (*10 ⁶)	23.98	23.98	10.17	13.81	0.91	0.24
R^2	0.28	0.40	0.21	0.36	0.26	0.35

Table reports coefficients from a regression of the performance gap (d_{nij}) on release year fixed effects. The unit of observation is an individual refueling transaction. Standard errors clustered by car type in parentheses. Columns vary as follows: (1) contains all data, (2) all the data with year and month, fuel type and vehicle model fixed effects, (3) only gasoline engines, (4) only diesel engines, (5) keeps observations only in the first year of driving (6) dependent variable is long difference between first and last observation for each driver instead of differences between visits.

Figure 3: Release year coefficients from fixed effect regressions

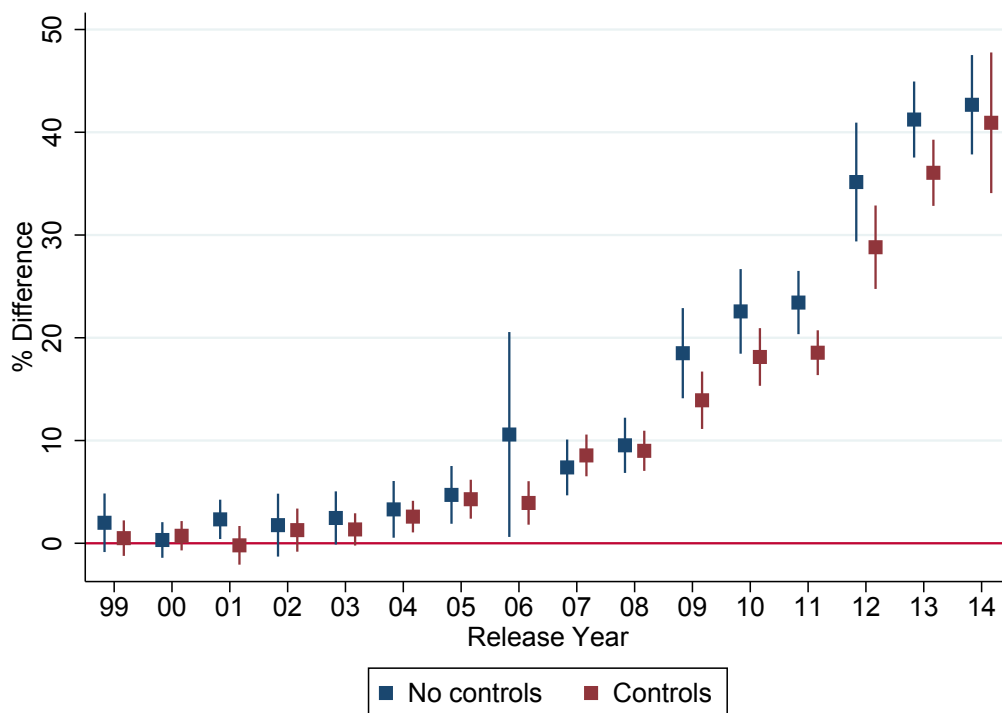


Figure plots coefficients from a regression of the performance gap (d_{nij}) on release year fixed effects. Coefficients correspond to regressions (1) and (2) from Table 1.

economy stay in our sample longer creating bias in the other direction. To control for this we keep only observations of d_{nij} that take place in the release year of the vehicle, so that we are capturing fuel consumption gaps only among the newest cars. Note that we are observing driving only from 2004 onwards, so we can only estimate on-road performance from vintages starting in 2004 in this specification. Column (5) reports those results. They show the same stark increase in the gap towards the end of the sample, which suggests that age effects are not driving our findings.

Finally, taking long differences, instead of differences between fuel station visits, might be more robust to bias from reporting errors and stockpiling. The drawback is that it does not allow us to control for time of driving effects in the regressions. Column (6) shows that our findings are robust: we get nearly identical estimates when we use long differences, taken between the last and first fuel station visit for each driver.

In sum, we conclude that the increase over time in the performance gap is robust to changes in fuel prices or congestion over time, to composition of the fleet, to rebound effects, to aging or selection over time, and to bias from stockpiling or reporting errors.

2.7 Heterogeneity in the performance gap does not rise over time

The welfare effects of gaming that we explore in the remainder of the paper potentially depend on whether there is significant heterogeneity in gaming within a release year across vehicle types.

To study potential heterogeneity in gaming at the vehicle level we need vehicle level estimates of the performance gap. In obtaining these estimates we want to take into account the underlying data quality for each vehicle controlling for the large variance in reported odometer readings, and the variation in the number of drivers and visits observed for different types of cars. Therefore, we follow the teacher value added literature, specifically [Chetty, Friedman, and Rockoff \(2014\)](#) and [Kane and Staiger \(2008\)](#), to estimate on-road fuel consumption r_j using precision weights and an empirical Bayes correction. We construct a precision-weighted mean of r_j and shrink it according to the reliability of the observations for j . Because of the measurement error in r_{nij} it is optimal, from a prediction standpoint, to use a biased but more precise estimate of each vehicle gap.¹⁰ The empirical Bayes procedure and accompanying variance decomposition is discussed in [Appendix A.2](#).

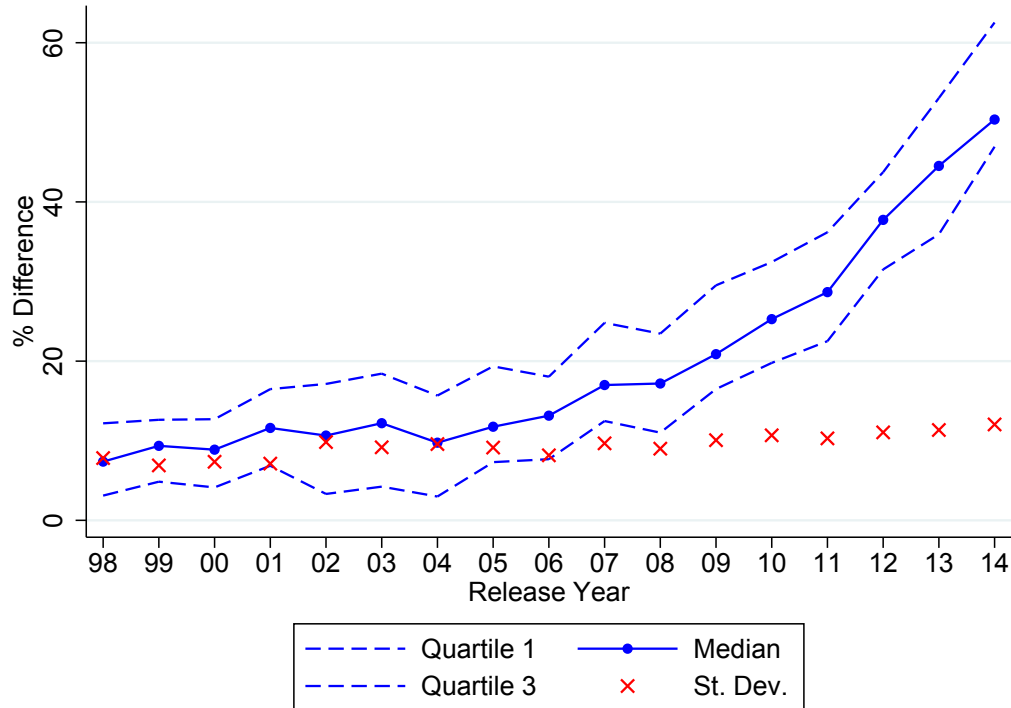
[Figure 4](#) plots the quartiles of the estimated Bayes corrected gaps for all release years as well as the standard deviation of the estimated fixed effects. Again we see a clear increase in the estimated gap over time, but perhaps surprisingly, this increase in the gap is not associated with a large change in the dispersion of the gap. If firms intensively game only some of their models, we would expect a rise in average gaming to correspond to a greater interquartile spread or larger standard deviation, across model types within a vintage. [Figure 4](#) shows, however, that these measures of dispersion did not rise over time. Instead, the data suggest that we should think about gaming in the EU market as an industry-wide shift, rather than a shift among select vehicles. This pattern is consistent with the underlying economics of gaming. Because the regulation is based on fleet averages, automakers have an incentive to game all of their vehicles, not just those with high consumption ratings. Moreover, to the extent that gaming consists largely of modifying test vehicles, the costs of gaming are largely a fixed cost, and it is plausible that techniques developed to game the test vehicle for one model line can be used for most other models.

2.8 Summary of reduced-form evidence

We conclude from this empirical section that there is very strong evidence of substantial gaming of fuel consumption ratings. Depending on our estimation method, the performance gap between on-road consumption and official test ratings varies between 10% to 15% before regulation. By 2012 this gap had increased to 30%, and all specifications estimate a mean gap of more than 45% in 2014. We have not found a single robustness check where this spectacular increase in the gap does not occur. Our findings regarding the time path of the performance gap is consistent with analysis performed by the International Council for Clean Transportation ([The International Council on](#)

¹⁰Note that this approach does not allow us to control for selection issues where drivers with different performance gaps select into different type of vehicles. Correlations between driving style and vehicle selection could influence our results, but again we emphasize our focus on changes in the performance gap over time, which mitigates these concerns.

Figure 4: Distribution of vehicle level fixed effects



Clean Transportation 2014, 2015).¹¹ Relative to those reports, we demonstrate robustness to the inclusion of various controls (most notably time period fixed effects), include standard errors, develop an empirical Bayes correction for sampling error, and calculate the distribution of the vehicle-level performance gap. Regarding the latter, this yields the striking and perhaps surprising finding that the gap is increasing for all vehicles and brands in the sample, so that the distribution of the gap is shifting, but the shape is relatively constant.

We see the decrease in the correlation between official and on-road ratings at a time when stakes on official ratings increase as a manifestation of Goodhart’s law. In the remainder of the paper, we interpret this time trend as evidence that automakers gamed official ratings in response to new regulations. We recognize that this evidence is suggestive, but we cannot think of any plausible alternative explanation.¹² While the EU commission has released an external evaluation of the emission standard, the official ratings are still the basis for the evaluation of the program and external communications about the program’s impact.¹³ Analysts who perform cost-benefit

¹¹The ICCT reports results from various sources including the data from Travelcard NV we obtained. The other sources include consumer self-reports uploaded to websites and consumer car testing magazines. The data from Travelcard are by far the most extensive in terms of number of vehicles and drivers included, but it is reassuring that all data in the ICCT reports point to very similar results.

¹²Tanaka (2017) shows quasi-experimental evidence on fuel consumption gaming in Japan. The Japanese standard specifies discrete notches, rather than a sales-weighted average as in our case, allowing for a comparison in gaps below and above a notch.

¹³See for example <http://ec.europa.eu/clima/policies/transport/vehicles/cars>, read on 09/13/2016.

analyses of the regulation may significantly mis-evaluate the program if they calculate savings based on official ratings.

We now turn our attention to our main economic question—who benefits from this gaming? Carmakers game the official rating to avoid stringent regulation, lowering compliance costs of the carmakers. Consumers will benefit from gaming to the extent these cost savings are passed through. But, these official ratings are also used by consumers when deciding what vehicle to buy. As a result, the manipulation of fuel consumption ratings alters the information value of the rating, which may distort consumer choice. Because of this, the incidence of the gaming in this market is ambiguous. We now set out to understand the incidence of gaming in response to corrective policy, when the policy measure has an information role in the market, beginning first with a model and then a calibrated market simulation.

3 A model of gaming

The goal of our model is to describe the impact of gaming on buyer welfare, taking into account how prices will depend on gaming, the presence of a policy, and the degree to which buyers are fooled by the gaming.

3.1 Setup

We model the market for a good that has a mutable attribute x that generates private value to buyers and also creates a social externality. In the automobile example, x is the vehicle’s fuel consumption rating. The full cost of the good is denoted $f = p + \beta x$, where p is the up-front purchase price and β is a coefficient that translates fuel consumption ratings into dollars.¹⁴ Note that x is a bad. We suppose that the good is sold by a monopoly seller, but the basic insights of our model do not depend on the form of competition. For simplicity, we assume that all attributes of a good are fixed other than x and p ; this can be interpreted as a short-run assumption.

Buyers do not observe x directly, but instead receive a message m from the seller (i.e., the official fuel consumption rating that appears on labels). Gaming occurs when the seller sends a message m that differs from x , where gaming g is defined as the difference between the true attribute and the message ($g = x - m$). Under full information, buyer demand D for the product depends on its various attributes (held fixed and thus suppressed in notation) and lifetime cost, which depends on x and p : $D(f) = D(p + \beta x)$. Buyers, however, observe m not x , so they use the observed signal to form beliefs, labeled \tilde{x} , where $\tilde{f} = p + \beta \tilde{x}$. We assume risk neutrality.

Buyer beliefs \tilde{x} are assumed to be a weighted average of the truth and the signal:

$$\tilde{x}_t = \alpha x + (1 - \alpha)(x - g) = x - (1 - \alpha)g.$$

¹⁴For a vehicle, β is the price of fuel per liter times the number of present discounted lifetime kilometers driven, while x is the L/100km rating.

This is a tractable form of beliefs that encompasses a variety of possibilities. When $\alpha = 1$, buyers can see through gaming completely. When $\alpha = 0$, buyers are completely fooled. For intermediate cases, $(1 - \alpha)$ represents the extent to which the buyer is fooled by gaming. The key for our purpose is that we want to allow the possibility that firms can fool buyers, to some degree. This specification is appealing in part because, for all of our results, we will be able to see how the effects differ when buyers are not fooled just by setting $\alpha = 1$.

Changing x or g incurs a cost. We assume that these costs are separable, and that changing x determines marginal cost, while changing g incurs a fixed cost. These assumptions about cost, which are convenient but not essential, are made in line with the market we study. Gaming is reported to be achieved largely through manipulation of the test vehicle, while real fuel consumption improvements require marginal costs.¹⁵

We write the marginal cost of production for a product as $c(x)$, which is decreasing (x is a bad) and convex ($c' < 0, c'' > 0$). The cost of gaming is denoted $h(g)$, which is increasing and convex ($h' > 0, h'' > 0$). The cost of gaming can include both real resource costs (e.g., engineering costs to game the test) and the risk of regulatory penalty.

Policy intervention is motivated by a negative externality associated with x , which we assume is linear and equal to ϕx . The regulator observes m , so policy must be based on m . In line with our empirical application, we focus on a fuel economy regulation that requires that the (reported) attribute be below a threshold: $\sigma: m = x - g \leq \sigma$. We will use λ to denote the shadow price of the regulation *per unit*; i.e., the constraint on the firm's profit function is written $\lambda \times (\sigma - x + g) \times D$. We later discuss an alternative policy that raises the price of fuel.

3.2 Summary of seller behavior

The monopoly seller chooses p , x and g to maximize profits, given the demand system, the degree of buyer sophistication, and policy interventions. The firm's Lagrangean is:

$$\mathcal{L} = (p - c(x) + \lambda(\sigma - x + g))D(p + \beta x - (1 - \alpha)\beta g) - h(g). \quad (3)$$

What values will the seller choose? Price will be chosen according to the standard optimal markup formula for a monopolist that equates marginal revenue and marginal cost per unit sold (given the profit-maximizing x and g).

The choice of x depends directly on policy. Absent policy ($\lambda = 0$), the seller will choose x so that $-c'(x) = \beta$, which is privately efficient.¹⁶ The seller lowers x until the marginal cost of doing

¹⁵We model a single-product setting for clarity. The main drawback to this is that it does not directly represent compliance that occurs through mix shifting—that is, shifting sales towards high efficiency models in order to meet a fleetwide average rating—which is the focus of our simulations. The impact of mix shifting will manifest as a set of price wedges for vehicles that operates similarly to the cost wedge modeled here. Our framework in general is applicable to any compliance strategy as long as compliance requires changes in pricing or marginal costs away from the firms' privately optimal choices, whereas our model is not designed to capture the buyer welfare effects of compliance methods that operate through fixed costs.

¹⁶The first-order condition for x is $0 = -c'(x)D + (p - c(x))\beta D'$. Substituting in the optimal markup from the first-order condition for price ($p - c(x) = -D/D'$) yields the result.

so equals the marginal value of increased energy efficiency to the buyer. The seller makes this privately efficient investment because any improvement in energy efficiency that is cost effective allows them to raise prices.

When there is a binding policy, the seller will choose x to solve $-c'(x) = \beta + \lambda$. This is the standard “internalization” of the externality suggested by the Pigouvian tradition. If there were perfect compliance and there was no distortion to quantities sold due to market power, then the first-best solution would be obtained by setting σ so that $\lambda = \phi$.

The seller’s choice of gaming will depend on the cost function $h(g)$. Specifically, the seller’s choice will satisfy $h'(g) = \lambda D - (p - c)(1 - \alpha)\beta D'$. This equates the marginal cost of gaming to the marginal benefit, which includes a regulatory benefit (λD) and a benefit from shifting out demand, which is equal to the markup $(p - c)$ times the shift in demand due to gaming $(1 - \alpha)\beta D'$.

3.3 Absent policy, gaming lowers buyer surplus

Our main interest is in establishing how gaming affects buyer welfare in this setting. We show how the firm alters its choice of p and x for any given level of g , which allows us to then characterize buyer surplus. The optimal degree of gaming will depend on the regulatory punishment function, so this exercise can be thought of as examining how shifts in the punishment function that alter the equilibrium g will influence buyer surplus.

In our analysis we focus on buyer surplus, rather than full social welfare, because we believe it is the incidence of gaming about which there is an interesting theoretical question. If the policy is set optimally, gaming obviously lowers total social welfare, once buyer surplus, seller surplus and the externality are taken into account. This is true by definition. Accordingly, we use the term “buyer surplus” to denote the private benefits, abstracting from the externality, rather than the more encompassing term of consumer surplus.

Our first proposition shows that, when there are no corrective policies, an increase in gaming hurts buyers. (All proofs are in Appendix A.3.)

Proposition 1. *In the absence of a binding policy ($\lambda = 0$), buyer surplus falls with the level of gaming. Specifically:*

$$\frac{dBS}{dg} \approx - \underbrace{\rho(1 - \alpha)\beta D}_{\text{price effect}} + \underbrace{(1 - \alpha)^2 \beta^2 D' g}_{\text{choice distortion}} \leq 0.$$

The effect of gaming on buyer surplus has two components. The first is a *price effect*. Gaming lowers *perceived* fuel costs of the product, which shifts out demand by the perceived reduction in full price, equal to $-(1 - \alpha)\beta$.¹⁷ The seller responds to this demand shift by raising prices. How much prices rise depends on the pass-through rate, which we denote ρ .¹⁸ The pass through rate

¹⁷Note that gaming does not affect the choice of x , which satisfies $-c'(x) = \beta$.

¹⁸For a monopoly seller, $\rho = (1 + (\epsilon_D - 1)/\epsilon_S + 1/\epsilon_{ms})^{-1}$, where ϵ is the elasticity of demand D , supply S and the inverse of the marginal surplus curve ms . Under another form of competition, the price effect will have the same formula but the pass-through rate will have a different functional form (Weyl and Fabinger 2013).

is typically used to describe how the burden of a tax is shared between buyers and sellers. An increase in gaming affects a shift in perceived demand that functions exactly like a tax to buyers, hence the analogy. This price effect is unambiguously bad for buyer surplus because the shift out in demand was a mistake—the buyer still has to pay the true operating cost of the vehicle, which is unchanged, but the seller has taken advantage of the mistake to raise prices.

The second effect on buyers of gaming is what we call the *choice distortion*. The buyer misperceives the true full cost f of the product due to gaming and thus purchases too much of the good given its true ownership cost. This misoptimization creates deadweight loss that is directly analogous to a Harberger triangle. Its width is the difference in demand, at the final price, induced by the gaming: $D(p + \beta x - (1 - \alpha)\beta g) - D(p + \beta x)$. Its height is the perceived gap in fuel cost induced by gaming: $(1 - \alpha)\beta g$. The term that appears in Proposition 1 is the derivative of this distortionary triangle. It is zero when $g = 0$; it is rising with the square of the distortionary wedge; and it is larger when demand is more elastic.

Both the price effect and the choice distortion unambiguously hurt buyers who are fooled by gaming. The perceived improvement in the product causes sellers to raise prices, and the buyers buy too much of the good. When buyers are fully sophisticated ($\alpha = 1$), gaming causes no change in buyer surplus; both terms go to zero. Both effects are larger when the degree of misperception ($1 - \alpha$) is larger. The price effect scales with the level of demand. The choice distortion scales with the slope of demand.

3.4 With a regulation, gaming can raise buyer surplus

Now suppose that there is a binding regulation ($\lambda > 0$). Proposition 2 shows that the regulation (potentially) flips the sign of the price effect, so that buyers may in fact benefit from gaming, even when they are fooled by it.

Proposition 2. *In the presence of a binding standard ($\lambda > 0$), a change in gaming affects buyer surplus as follows:*

$$\frac{dBS}{dg} \approx \underbrace{(-\rho(c' - \alpha\beta) - \beta)D}_{\text{price effect}} + \underbrace{(1 - \alpha)^2\beta^2 D'g}_{\text{choice distortion}} .$$

As in the case with no regulation, the effect of gaming on buyer surplus includes a price effect and a choice distortion. The choice distortion formula is not affected by the policy, it is the same as in Proposition 1 and is unambiguously negative (lowers buyer surplus).

The price effect, however, is significantly altered by the externality-correcting regulation and may be positive. Gaming causes three effects that determine the full price ($p + \beta x$). First, the seller will use an increase in g to reduce x under a binding standard (i.e., $dx = dg$ because $\sigma = x - g$ when the standard binds). This raises the true full cost of ownership by $-\beta$ per unit sold. Second, true cost of production falls by amount $c'(x)$. Some of this cost savings—equal to $\rho c'$ —will be passed through to buyers via a reduction in p . Third, buyers will perceive a change in operating cost. This

shifts in demand, and the seller will adjust p accordingly in reaction to the perceived operating cost increase. Specifically, with $dx = dg$, the buyer will perceive an operating cost change of $\alpha\beta > 0$, of which $\rho\alpha\beta$ will be finally born by buyers.

The sign of the full price effect is ambiguous. The term $-\rho(c' - \alpha\beta)$ is positive (it raises buyer surplus), but it may or may not be larger than the rise in true costs (equal to β), which lowers buyer surplus. When the price effect is positive (buyers gain), it may be larger than the negative choice distortion, so that gaming benefits buyers. Intuitively, the benefit to buyers of gaming comes because the seller experiences a cost reduction when they avoid the costly regulation. By definition of a binding standard, the standard forces x to move away from the privately efficient level, so gaming provides a private welfare gain by allowing the seller to move back towards that level. Even if buyers are unaware of the firm's gaming, they will enjoy some of the cost savings due to standard pass through forces.

Several key comparative statics that we return to in our simulation analysis are apparent from Proposition 2. Empirically, we are particularly interested in establishing conditions under which buyers are likely to benefit from gaming. Factors that influence this outcome include the following:

1. More sophisticated (higher α) buyers gain more (or lose less) from gaming.
2. Tighter standards (higher λ) imply larger buyer gains (or smaller losses) from price effects.
3. More gaming (higher g) implies smaller gains (or larger losses).

On the first point, more sophisticated buyers experience smaller choice distortions, and they will also get more favorable price effects.

On the second point, for a given amount of gaming, price effects get more positive (favorable to buyers) under a tighter standard because gaming creates a larger cost savings. For a sufficiently tight standard, the price effect will always be positive. This is true even if buyers are fully unaware of gaming.

On the third point, for a given standard, the shadow price of the regulation is falling as gaming increases by concavity of the profit function. In contrast, the choice distortion is rising in gaming, because the total choice distortion is a function of the square of g .

To recap, under a binding standard, gaming induces a price effect of ambiguous sign, and a choice distortion that lowers buyer surplus. When choice distortions are sufficiently small, positive price effects can dominate, so that increases in gaming raise buyer surplus, even when the buyers are fooled by the gaming. In the final part of the paper, we conduct calibrated simulations of the European vehicle market to show that it is indeed likely that car buyers experienced net benefits from gaming, even if they were completely fooled by it.

3.5 With a fuel tax, gaming lowers buyer surplus

Above we considered a fuel economy regulation, which is the most important carbon policy for automobiles in the EU, as well as the United States, China and Japan. A regulation of this type is

theoretically identical to a revenue-neutral set of product taxes (sometimes called a feebate) that is linear in the fuel consumption rating. Thus, our results can be interpreted as informing the analysis of either a feebate or a regulation. An alternative policy is a tax on motor fuels. In our model, a fuel tax simply raises β , which is the term that translates fuel consumption ratings into present-discounted value, so it is straightforward to show how a fuel tax will impact buyer surplus.

Proposition 3 shows that under a fuel tax, denoted as τ , that raises the lifetime operating cost of the vehicle to $f = p + (\beta + \tau)x$, an increase in gaming unambiguously lowers buyer surplus.

Proposition 3. *In the presence of a fuel tax but not a binding standard ($\lambda = 0, \tau > 0$), buyer surplus falls with the level of gaming. Specifically:*

$$\frac{dBS}{dg} \approx - \underbrace{\rho(1-\alpha)(\beta+\tau)D}_{\text{price effect}} + \underbrace{(1-\alpha)^2(\beta+\tau)^2 D'g}_{\text{choice distortion}} \leq 0.$$

When the planner uses the fuel price instrument instead of the regulation, the benefits of gaming no longer accrue to buyers in the same way as under a standard. Gaming does not help the buyer avoid the fuel tax τ , so it provides them no benefits. Unlike the case of a binding standard, gaming under the fuel tax operates just like gaming in the case of no regulation. The only difference is that β has been raised by policy.

Normally, the statutory incidence of a tax is irrelevant to its ultimate economic incidence; i.e., it does not matter if a retail sales tax is levied on buyers or sellers. But, this irrelevance need not hold when there are opportunities for avoiding or evading the tax (Slemrod 2008, Kopczuk et al. 2016). Our result here is a manifestation of this same result. In our model, a regulation σ creates a shadow price λ that corrects the externality associated with x . A tax τ does exactly the same thing. If there were no evasion, the two policies would be identical whenever $\tau = \lambda$, but gaming offers an opportunity to avoid a regulation but not a fuel tax, which breaks the symmetry and provides an enforcement rationale for preferring the tax over a regulation.

3.6 The environmental consequences of gaming

Theoretically, we are focused on the question of the incidence of gaming, but it is worth returning now to a discussion of the externality. The environmental impact of gaming depends upon the degree to which the planner adjusts policy to account for gaming, as well as the magnitude of what we call the mechanical and market size effects. For a given standard, gaming has a mechanical effect on pollution—gaming substitutes one for one for emissions reductions. Less obvious is the potential market size effect. If buyers are fooled by gaming, then gaming may lower the perceived price and the total market size will increase, which also increases pollution.¹⁹

¹⁹Note that, in the presence of market power, pollution increases due to an expansion of the market may not be welfare decreasing, in the spirit of Buchanan (1969).

This market size effect means that, on net, the introduction of a policy could *increase* pollution. For example, let x^b be the equilibrium attribute absent policy. A planner introduces a binding standard $\sigma < x^b$. Suppose that the firm complies exclusively through gaming. If buyers are partially naïve, they will perceive an improvement in the product. The net effect will be a perceived price reduction compared to the baseline, which leads to larger total demand than in the pre-policy baseline. Because the true attribute is unchanged from the baseline, the market expansion implies an unambiguous total increase in pollution due, caused by the policy.

Given the extensive degree of gaming in the European vehicle market, this scenario is not necessarily far fetched. Moreover, it is highly concerning for international climate negotiations that are structured like the Paris Accord, under which there may be considerable pressure for individual nations to deliver carbon reductions “on paper” that they do not actually wish to enforce. If lax enforcement leads to inflate energy efficiency statistics that partially fool buyers into perceiving cheaper durable goods, then this could act to raise total emissions.

Suppose instead that a regulator really does wish to lower emissions and is aware of gaming. If both the planner and buyers are aware of the degree of gaming, then the planner can simply ratchet down the standard to achieve the true desired level of emissions, accounting for gaming. In this case, gaming will still occur, but there is no choice distortion and the only inefficiency comes from resources spent to game, embodied in $h(g)$. To the extent that these are real resource costs, this is a social waste, but much of $h(g)$ may represent regulatory risk.

More interesting then is the case when the planner is aware of gaming, but buyers have less than perfect sophistication ($\alpha < 1$). In this case, the regulator can ratchet down the standard in order to achieve the desired level of pollution. But, following the logic of Goodhart’s Law, where tighter standards induce more gaming, tighter policies will induce a larger choice distortion. The choice distortion raises the social marginal cost of abating pollution, which will lead the regulator to attenuate the second-best policy away from the Pigouvian benchmark, under which the standard is tightened until λ equals marginal damages per unit of x .

4 Calibration of the incidence of gaming

To quantify the effects of gaming, we estimate a discrete choice model for new vehicles in order to obtain the taste parameters of consumers. We show that there is a direct analog to the choice distortion and price effects identified by our theory that can be decomposed from the standard log-sum welfare formula for discrete choice models. Given this decomposition and the estimated taste parameters, we calibrate outcomes under various assumptions about consumer awareness, policy stringency, and the degree of gaming. This allows us to show comparative statics that illustrate how the incidence of gaming depends on key parameters. We use parameters informed by the actual market context, but we stress that we show results for a range of parameters to illustrate how the incidence of gaming depends on these values.

Our theory describes a monopolist, but our simulation is broader and considers a multi-firm,

multi-product setting. The main insights regarding the price effects are similar in either case, but with multiple products the choice distortion comes both from switches between products and from changes in the overall size of the market; that is, from the outside good. The final step of our simulation relates to competition and gaming. We show that regulation increases pressure on firms to game in this setting.

4.1 Setup and parameter choice

Buyer preferences: We begin by modeling the choice of a consumer making a discrete choice about which vehicle to buy. Each consumer i chooses the vehicle that maximizes her indirect utility, which we write as:

$$u_{ij} = \Delta_j \gamma_i - \eta p_j + \beta_i (\kappa_j * x_j) + \xi_j + \varepsilon_{ij},$$

where Δ_j is a vector of vehicle characteristics, p_j is price divided by income and $\kappa_j * x_j$ is the operating cost of the vehicle for driver i , measured as fuel consumption transformed into euros per kilometer using fuel prices κ_j .²⁰ Utility from the outside good (not buying a new vehicle) is normalized to zero, $u_{i0} = \varepsilon_{i0}$. We estimate a random coefficient logit model and assume that β_i and γ_i are independently normally distributed. We estimate the mean and standard deviation of β_i and γ_i . All remaining consumer heterogeneity is contained in the additive idiosyncratic error term ε_{ij} . The ξ_j term represents the value of product attributes unobserved by the researcher but observed by firms and consumers. Following [Berry, Levinsohn, and Pakes \(1995\)](#), we integrate out ε_{ij} to construct choice probabilities. After inverting the choice probabilities to obtain ξ_j , we use a GMM estimator to estimate the vector of taste parameters.

Our dataset, a panel containing sales, prices and characteristics for all new vehicle sales in seven European countries between 1998-2011 allows us to estimate a rich demand model with flexible cross price elasticities. We only use data from before 2008 to estimate demand, so that our estimates come from a period in which the performance gap was stable. We thus assume variation in x_j is informative about actual fuel cost differences. Note that we assume that consumers have on average a correct belief about future fuel costs before gaming is introduced.²¹ Estimated taste parameters and standard errors are reported in appendix Table A.5. Price and fuel costs have the expected negative effect on utility. The mean of the estimated own price elasticities is -5.45, in line with the previous literature. In the appendix we detail the estimation procedure, which is very similar to the model estimated in [Reynaert \(2017\)](#).

Profits: In the empirical analysis we consider the impact of gaming with and without policy. The Lagrangean that the firm faces is equivalent to the one defined in (3) but firms optimize profits

²⁰Notice there is a slight change of notation relative to the theory: here β captures the willingness to pay for fuel costs while in the theory β captures both the willingness to pay for fuel cost as well as the rescaling of fuel consumption into money. In the demand estimation we rescale fuel costs using observed fuel prices κ_j .

²¹If there were undervaluation of fuel costs, gaming could mitigate externality costs as described in [Allcott and Wozny \(2014\)](#), and as modeled in [Allcott and Taubinsky \(2015\)](#).

taking the products in their portfolio as given. Firms thus maximize profits by choosing price (for the vehicles they produce), taking levels of x as given:

$$\begin{aligned} & \max_p \sum_{j \in F} [(p_j - c_j)q_j(p, g)] \\ \text{s.t. } & \sum_{j \in F} q_j \cdot (x_j - g_j) / \sum_{j \in F} q_j = \sigma \end{aligned}$$

in which F is the set of products a firm produces and q_j is quantity demanded, which depends on the full vectors of prices and gaming. Marginal costs c_j are not observed in the data. Because we estimate demand in a period where there is no regulation we are able to back out marginal costs from the unconstrained first order conditions. Given these marginal costs and estimated preferences we solve for new equilibrium prices with different levels of g and σ as we explain now.

The policy, as stated in the constraint, is a requirement for each firm to decrease its sales-weighted emissions and mimics the EU policy.²² When there is a binding policy, selling polluting products makes it harder for the firm to comply. Firms thus must adjust their price schedules in order to reach the target. When we introduce gaming, these changes will be lowered because gaming makes the constraint less binding (it lowers the policy wedge λ). Gaming also affects demand and optimal prices directly when consumers are at least partly fooled because the demand shifts out in response to perceived higher quality.

Policy target and abatement cost: We model the policy as a requirement for each firm to decrease its sales-weighted emissions by 5%. This 5% is a limited decrease but it mimics a situation in which firms can potentially comply without changing characteristics of vehicles. We model honest compliance as sales mixing, such that firms comply by changing prices to shift their sales to efficient vehicles in order to become compliant. The median price change is substantial, at €1860. This policy would be optimal, given a sales mixing compliance strategy, if the externality from a ton of carbon was €286. In the simulations, we also model more and less stringent policies that mandate 3% and 10% reductions.

Our results are representative of any setting where a firm is distorted from its optimal pricing because of a policy wedge. The same mechanisms are at play if there were emissions-based vehicle taxes (paid at the moment of purchase), feebates or technology adoption requirements. In the longer run, and when policy targets become more demanding, we expect firms to resort to technology adoption or downsizing of the fleet to comply. These mechanisms not only increase the marginal costs of vehicles, but also come with fixed costs. See [Reynaert \(2017\)](#) for a comparison of the welfare effects of different abatement strategies. Here, we focus on understanding the determinants of the incidence of gaming with sales-mixing as the benchmark for honest compliance. When we lower

²²We use the actual formula used by the EU in its emission standard, so that the policy is based on vehicle weight. For each firm we compute a sales weighted average emission rate, with emissions for heavier (lighter) cars receiving a bonus (penalty) in the weighted sum. Specifically, the emissions rating of each vehicle j in the weighted sum receives a penalty of $0.046 * (weight_j - 1370)$ where weight is measured in kilograms.

the stringency of the policy, the policy wedge decreases. This mimics the setting where compliance becomes cheaper for firms in the longer run.

Level of gaming: We model 5% gaming as a way to describe what the effects are of gaming when it completely offsets the policy. When we introduce gaming, firms that game can meet the policy requirement without a sales shift. In our simulations we also show what happens to buyer welfare when gaming is lower or higher than the policy target. Consistent with media coverage of the EU case, our procedure assumes that gaming is a fixed cost: it does not shift the marginal costs of the products. Buyers will perceive fuel costs as lower so that demand shifts out and prices increase. The market equilibrium will thus change to the extent that consumer demand, and hence prices and quantities, shift in response to gaming. Our procedure estimates the equilibrium prices and market shares for different levels of gaming, but it does not attempt to describe the optimal amount of gaming (this would require information on the costs of gaming). Rather, our goal is to use our demand system estimates to determine whether or not the price effects from gaming could plausibly dominate the choice distortion, leading to increases in consumer welfare even when consumers are naïve, and how this result hinges on the existence of a policy.

Consumer awareness: A crucial parameter is consumer awareness α . When consumers are fully aware of gaming our model collapses to a simple setting where regulatory avoidance benefits both buyers and sellers. In our empirical setting it is unlikely that consumers were aware of the full extent of gaming, but it is also unlikely that consumers did not update their beliefs over time.

A recent literature (see [Busse, Knittel, and Zettelmeyer 2013](#), [Allcott and Wozny 2014](#), [Sallee, West, and Fan 2016](#), [Grigolon, Reynaert, and Verboven Forthcoming](#)) has shown that buyers do take fuel consumption ratings into account when purchasing a vehicle. In a field experiment [Allcott and Knittel \(2017\)](#) show that treating buyers with detailed information about fuel consumption has no impact on their choice. In sum, this research suggests that consumers are well informed about fuel consumption and value savings in future consumption when choosing between alternative engines. However, in our setting buyers are confronted with very rapid declines in the information content of reported fuel consumption due to gaming.

Might consumers have been informed by the media? A news search suggests that the first reports about gaming appeared in the fall of 2013 in Germany. In other countries we find that the first newspaper articles appear only in 2014, by which time we estimate a fuel consumption gap in excess of 40%.²³ In 2015 there is a major news episode related to the Volkswagen crisis. This suggests that there was not EU-wide media coverage of the performance gap until 2015.

One other source of information are websites where consumers upload information about their experienced fuel consumption. Several such website exist during this time period. We note, however, that our empirical results demonstrate that there is significant heterogeneity across trips and drivers

²³In particular, we perform a Google news search with the words “ICCT” and “gap” in German, English and Dutch. The ICCT reports were discussed in the German media in 2013 and 2014, but we find no references in other languages, except for two automobile websites in English. Other word combinations gave no additional results.

in experienced fuel consumption due to driving conditions. This makes it difficult for consumers to distinguish legitimate variation from gaming without a large sample. Thus, we conclude that consumers might have had some information, but the available information was far from perfect.

We have tried to directly estimate consumer awareness of gaming with our data, but did not find statistically meaningful results. The challenge is that our estimate of on-road consumption is a function of gaming and selection, and the lack of heterogeneity in gaming across models within a year (as shown in Figure 4) makes it difficult to isolate meaningful variation between vehicles that can be used to explain market shares.²⁴

In what follows we show how the price effect and choice distortion vary when consumers range from completely naïve to fully sophisticated (α ranges from 0 to 1). We argue that in the long term buyers will update their beliefs, but in the short term consumers may be distorted in their choices by gaming.

Buyer surplus To compute buyer welfare when consumers are affected by gaming ($\alpha < 1$, in terms of our theory) we follow [Dubois, Griffith, and O’Connell \(2018\)](#), who describe the welfare effects of persuasive advertising. As described in Section 3, gaming changes prices, as firms will exploit higher demand from lower perceived fuel costs. We label the equilibrium price without gaming p_j^0 , and the prices with gaming as p_j^1 . The misperception of gamed fuel costs also distorts consumer choice. To separate the two effects we make a distinction between decision (at the moment of purchase) and experience utility (at the moment of utilization).²⁵ When there is gaming, a naïve consumer will perceive fuel costs, following her belief, as $\tilde{x}_j = x_j - g_j$, in which g_j is the increase in gaming. The consumer will make her choice based on \tilde{x}_j and will perceive her decision as yielding a utility of:

$$\tilde{V}_{ij}(g, p^1) = \Delta_j \gamma_i - \eta p_j^1 + \beta_i \tilde{x}_j + \xi_j + \varepsilon_{ij}.$$

After purchasing the vehicle, true fuel costs are revealed, and the consumer has experience utility:

$$V_{ij}(g, p^1) = \Delta_j \gamma_i - \eta p_j^1 + \beta_i x_j + \xi_j + \varepsilon_{ij}.$$

The difference between decision and experience utility is the optimization error, which in this case is the value of the additional fuel costs for the consumer $\beta_i(x_j - \tilde{x}_j) = \beta_i g_j$. Buyer surplus with

²⁴In particular, we did the following. We regressed our vehicle level gap estimates on a set of explanatory variables and predicted gaps for vehicles in the choice set for which we have no on-road estimates. We then tested if differences in gaps explain differences in market shares by introducing the estimated gaps in our discrete choice framework. This resulted in very noisy estimates with a wide confidence interval for the effect of the gap on sales. In a second stage, we tried to control for selection, using the assumption that gaming is absent prior to 2007. This allows us to decompose the gaps post 2007 into gaps from driver selection and gaming. When we introduce this measure in the discrete choice framework we find no statistically (or economically) meaningful estimates. One key limitation is that our sales data end in 2011, which excludes the period of greatest interest. A related exercise that tests for changes in pricing due to gaming similarly yielded only noisy results. We thus cannot make any statistical statement about consumer awareness of gaming based on our data.

²⁵This is conceptually the same as the choice distortion described in [Allcott \(2013\)](#) and [Sallee \(2014\)](#) who study misperception of fuel costs, but our approach here is more general because we allow firms to change prices in response to gaming. [Leggett \(2002\)](#) also models a similar distortion in a discrete choice setting.

gaming can then be written as:

$$\begin{aligned}\widetilde{W}_i(g, p^1) &= \mathbb{E}_\varepsilon[\widetilde{V}_{ij}] - \mathbb{E}_\varepsilon[\beta g_j] \\ &= W_i(g, p^1) - \sum_j [s_{ij}\beta g_j],\end{aligned}$$

where s_{ij} are the choice probabilities obtained from maximizing the decision utility. We compute $W_i(g, p^1)$ by applying the log-sum formula of [Small and Rosen \(1981\)](#). We can then decompose the change from the equilibrium with honesty ($0, p^0$) to the equilibrium with gaming (g, p^1) into a price effect and choice distortion:

$$W_i(0, p^0) - \widetilde{W}_i(g, p^1) = \underbrace{W_i(0, p^1) - \widetilde{W}_i(g, p^1)}_{\text{Choice Distortion}} + \underbrace{W_i(0, p^0) - W_i(0, p^1)}_{\text{Price Effect}}.$$

When there is no policy the reference prices p^0 are the observed prices in the market. When we introduce gaming without policy, prices p^1 will be higher than p^0 because of market power. When we introduce the policy, the reference prices p^0 become the prices obtained from honest compliance to the policy. These prices will differ from the no-policy reference prices because of the tax wedge that the policy introduces ($\lambda > 0$). When we introduce gaming, the prices p^1 will change for two reasons: gaming lowers the tax wedge and firms with market power increase prices when gaming. In contrast, the choice distortion will be similar with and without policy as it is not a function of the tax wedge.

Note that consumer choice is distorted by gaming even when all products are gamed an identical amount. Intuitively, a uniform reduction in fuel costs is akin to a reduction in the coefficient on fuel costs in the utility function. Thus, when all firms game and perceived fuel costs drop, consumers will, to some degree, reorder their favored products because they effectively put more weight on other vehicle characteristics.²⁶

Procedure: In all simulations, we start with a base market, which is the observed market in the Netherlands in 2007. That is, we use several markets over several years to estimate the demand system, but we need to choose one particular market as our baseline from which to calculate changes. We compute consumer surplus from this observed market and our parameter estimates. To simulate gaming, we first lower the perceived fuel cost of the products made by firms that game. Second, with these new fuel costs, we solve for demand and prices using the first-order conditions for profit maximization. Third, we compute changes in welfare relative to the reference market with prices p^0 . The obtained changes in welfare are the changes in yearly utility, profits and emissions from new vehicle sales.

²⁶Note that we limit ourselves to reporting changes in consumer welfare from purchasing a new vehicle. We do not consider effects on the second hand market or on the amount that consumers drive. Consumers could potentially react to gaming by changing the amount they choose to drive in the purchased vehicle, a potential effect that we also abstract from.

4.2 Incidence of gaming

In this section we simulate the effect of gaming on buyer surplus and pollution. We start from a base scenario in which we impose a policy of 5% that is responded to by 5% gaming. We focus first on the role of consumer awareness before discussing policy stringency and the level of gaming.

Awareness: In Figure 5 we plot buyer welfare changes from 5% gaming when there is no binding policy and when there is a 5% policy. When there is no policy both the choice distortion and the price effect are negative: gaming hurts buyers. Both the choice distortion and the price effect go to zero when consumer awareness α increases from 0 to 1. When we introduce the policy, the choice distortion remains very similar and again goes to zero when consumer awareness increases. The price effect, however, switches sign and becomes positive and an order of magnitude larger than the choice distortion. This causes overall buyer surplus to *increase* with gaming. The intuition for the switch in the price effect is cost avoidance of regulatory compliance that is passed through from seller to buyer. Notice that this positive price effect is very similar regardless of consumer awareness; pass through occurs whether or not consumers are fooled. The negative effects from gaming via choice distortion increase when consumers are fooled, but even when consumers are fully unaware of the gaming, the price effect strongly dominates.

Table 2 shows the effect of gaming on buyer surplus and pollution when consumers are unaware ($\alpha=0$). Column I is the no policy case, where we show the decrease in buyer surplus decomposed into both a negative choice distortion and a negative price effect. Column II shows the expected switch in sign of the price effect when a policy is in place. It shows that the price effect is an order of magnitude larger than the choice distortion, so that buyer surplus increases with gaming under a policy.

Table 2 also shows the effect of gaming on pollution. The first line shows the percentage-point change in total sales, which captures how the size of the market changes as a result of gaming. The second line shows the percentage change in average fuel consumption, which captures to what extent buyers substitute toward more polluting vehicles as a result of gaming. When there is no policy (Column I), we see that gaming increases the market size by 3% and that consumers buy vehicles that are slightly more polluting (0.44%). When a policy is in place (Column II), these effects increase considerably to a 4.5% increase in market size and a 5.8% increase in average fuel consumption. The reason is that the reference point is honest compliance to the policy with a target of 5% reduction in average fuel consumption. Not only does gaming undo the target (the mechanical effect), but consumers also buy more vehicles and shift toward more polluting vehicles when they are fooled by the gaming (the market size effect).

Proposition 2 in our theory showed that it was possible that consumers benefit from gaming, even when they are fooled by it. Our simulation results suggest that this is indeed the case in the European auto market. In line with Proposition 1, we find that consumers unambiguously lose from gaming when there is no policy. But, this result is overturned with the introduction of a policy. Consumers no longer benefit from honesty. Moreover, when choice distortions are significant,

Figure 5: Consumer awareness and incidence of gaming

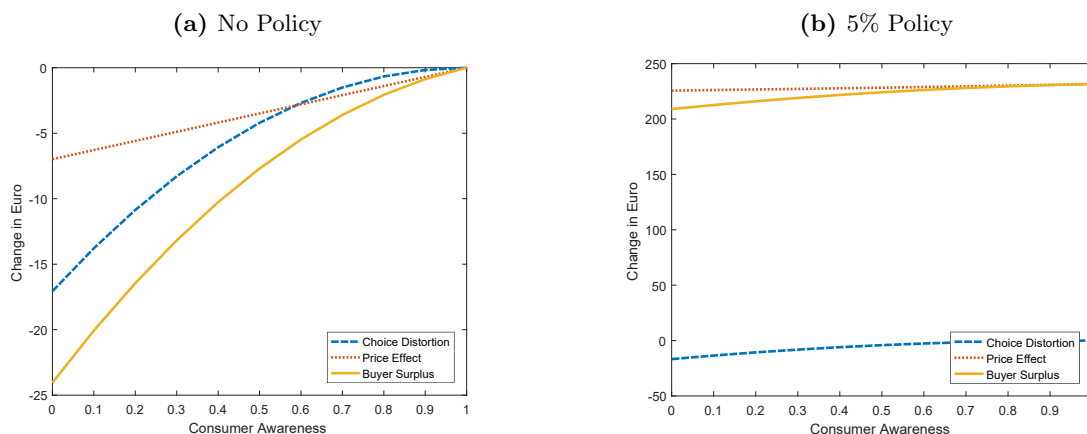


Figure plots buyer surplus, choice distortion and price effect in Euro when consumers range from unaware ($\alpha = 0$) to aware ($\alpha = 1$). In panel (a) we model 5% gaming and no policy, in panel (b) we model 5% gaming with a 5% policy.

emissions actually *rise* in response to the policy, due to a combination of weak enforcement and significant shifts in the market due to choice distortions. We believe this raises an important concern for policies, like current international climate negotiations, that rely on voluntary on-paper emissions pledges rather than effective and credible measurement.

Note, that in all scenarios in Table 2 consumers are unsophisticated and thus fooled by gaming. As shown in Figure 5, this is the scenario in which negative effects of gaming are largest for consumers. Evaluating the effects of gaming at $\alpha = 0$ is thus conservative against finding positive effects on buyer surplus.

Stringency of policy and level of gaming: In the second panel of Table 2 we vary the stringency of the policy from 3% to 5% to 10%, while keeping the degree of gaming constant at 5%. In all cases, the price effect is positive and larger than the choice distortion. The price effect is only marginally larger than the choice distortion for the 3% target, but the price effect outweighs the choice distortion by hundreds of euros when the standard becomes more stringent. This is in line with the theory; the price effect is increasing with the stringency of the regulation. This is relevant for assessing the EU standard, which is arguably the most demanding fuel consumption standard in the world. Because it is very stringent, it is likely that price effects outweigh choice distortions and lead to net consumer surplus gains.

Figure 6 shows buyer surplus and its decomposition when we increase the level of gaming from 0% to 10% in response to a 5% policy. The choice distortion hurts consumers more and more as we increase gaming. The price effect is a parabola because of two off-setting effects. The curve is steepest with the first units of gaming as these soften the burden of the policy the most. With the following units of gaming, up until 5%, more and more of the policy is avoided but this gives smaller incremental increases. At the same time higher gaming increases prices because of market

Table 2: Effects of 5% Gaming on Buyer Surplus and Pollution

Base and Alternative Scenarios				
Scenario:	I	II	III	IV
Policy:	0%	5%	5%	5%
Buyer Surplus (€per household)				
Total BS	-24.78	208.27	186.29	627.6
Choice Dist.	-17.09	-16.72	-30.74	-8.6
Price Effect	-7.7	224.99	217.04	636.2
Pollution (% change)				
Sales	3.05	4.58	4.52	0.51
Emissions	0.44	5.82	6.19	5.68
Increasing Policy Stringency				
Scenario:	V	VI	VII	VIII
Policy:	0%	3%	5%	10%
Buyer Surplus (€per household)				
Total BS	-24.78	33.19	208.27	608.74
Choice Dist.	-17.09	-17.17	-16.72	-13.47
Price Effect	-7.7	50.69	224.99	622.22
Pollution (% change)				
Sales	3.05	1.83	4.58	8.25
Emissions	0.44	5.99	5.82	5.48

Table gives changes in buyer surplus and pollution from gaming for unsophisticated consumers ($\alpha = 0$). Top panel: (I) 5% gaming when no policy is in place; (II) 5% gaming when a 5% policy is in place; (III) uniformly distributed 0-10% gaming when a 5% policy is in place; (VI) 5% gaming when a 5% policy is in place and when outside good share is 10% instead of 61%. Bottom panel increases the policy stringency from 0% to 10% while keeping gaming constant at 5%. (Note that V=I and VII=II.)

Figure 6: Level of Gaming

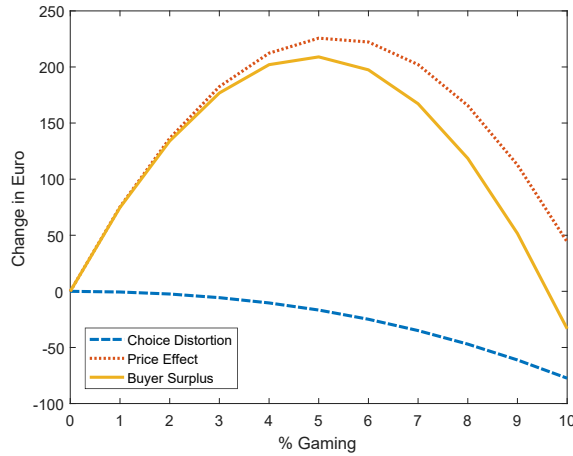


Figure plots buyer surplus, choice distortion and price effect in Euro when gaming ranges from 0% to 10% in response to a 5% policy.

power. This effect becomes dominant once the policy is fully avoided and thus the price effect decreases once more gaming is done than required to comply with the policy. The combination of the decreasing choice distortion and parabolic price effect gives an overall buyer surplus that is also parabolic. Note that gaming has positive effects on buyer surplus until 9%, almost twice the level required to comply with policy.

In appendix Table A.6 we also show results when the policy is a fuel tax instead of a standard.²⁷ This confirms that gaming with a fuel tax is harmful to buyers, consistent with Proposition 3. The table shows results from fuel taxes that were found to produce the same change in fuel consumption as the standards. We find that gaming decreases buyer surplus in very similar magnitudes to our base specification when firms game in the absence of a policy. Buyers will not benefit from gaming in response to a fuel tax because gaming will distort their choice; the price effect will lead them to pay higher, not lower, prices for the vehicle; and gaming does nothing to help them avoid the fuel tax. When they show up at the pump, they pay the tax per liter actually consumed, regardless of the vehicle’s official rating.

Heterogeneity in Gaming and Outside Good: The choice distortion drives the negative effects of gaming on buyer surplus and pollution. Because fuel consumption ratings are shaded by the gaming, buyers purchase the wrong amount and the wrong type of vehicle. In Table 2 columns I-II and V-VIII, we see large increases in sales (between 1.8% and 8% points) and limited increases in average fuel consumption per vehicle (between 0.4%-0.9% on top of the mechanical 5%). Gaming thus leads to considerable changes in the size of the market and in the type of products purchased. The size of these two effects, however, depends partly on two underlying assumptions in

²⁷We also show results for a flat emission standard that is not attribute based. This flat standard is less demanding for most firms and so the price effects are smaller, but still larger than the choice distortion

our simulations: the lack of heterogeneity in gaming and the amount of substitution to the outside good.

In Column III of Table 2 we relax the assumption of homogenous 5% gaming on all products and instead draw gaming from the uniform distribution $[0\%, 10\%]$, so that mean gaming is 5% as in the main scenario. Now gaming not only reduces the importance of fuel consumption to choice relative to other attributes, but also scrambles the ranking of preferred products more directly by providing a bigger improvement to some products than to others. This has two implications relative to the homogeneous 5% gaming in Column II. First, we see that the choice distortion almost doubles from €17 to €31. Because gaming is heterogeneous, buyers buy the wrong type of cars in larger amounts and end up further from their preferred vehicle. Nevertheless, despite the doubling of the choice distortion, the price effect is still much larger as the pass through of regulatory avoidance is not affected: on average products have 5% lower official fuel consumption so that the regulation is still avoided. Second, the environmental damage due to gaming is larger and average emissions now increase by 1.2. This is because consumers substitute more toward high quality and high pollution vehicles that were gamed relatively more.

In Column IV we change the size of the outside good in our estimation and in our simulations. The estimated demand parameters do not change significantly because we include market fixed effects, but the results of the simulation do change noticeably. In Column II we obtained numbers from a model in which we assume that one out of seven households are potentially interested in purchasing a new car each year such that the outside good share is 61%. This assumption determines the price elasticity of the industry. When we lower the price elasticity of industry by decreasing the outside good share to 10%, we find significant changes. First, the positive price effects increase significantly because buyers do not stop purchasing polluting vehicles even at much higher prices. The price changes needed for firms to comply thus increase and the benefits of avoiding the regulation are larger. Second, the choice distortion shrinks. Because the price elasticity of industry is much lower now, buyers will not change their quantity drastically in response to the gaming. This is also important for pollution. The size of the market increases by 0.5% instead of 4.5% points, though buyers still substitute to more polluting vehicles on average. In sum, the positive price effect increases and the choice distortion decreases when the outside good share shrinks.

4.3 Competition and gaming

So far we have assumed that the whole industry games, in line with the empirical evidence shown in Section 2.5. Here, we study what happens if not all firms game. From the point of view of a single firm, gaming confers a competitive advantage. This is much the same intuition as for the monopolist; the residual demand curve faced by the firm can be shifted out by gaming. Conversely, when other firms are gaming, an honest firm would be at a competitive disadvantage. Thus, we expect competitive pressure to induce gaming, as was found in Becker and Milbourn (2011) and Bennett, Pierce, Snyder, and Toffel (2013).

In Table 3 we compute market outcomes with asymmetry in gaming. In the left panel only one

Table 3: Market changes from asymmetric gaming

	Gaming Alone		Not Gaming Along	
	No Policy	Policy	No Policy	Policy
	Consumer Surplus (€per household)			
Total	-3.83	15.79	-24.31	189.94
Choice Distortion	-3.17	-3.24	-17.23	-16.79
Price Effect	-0.66	19.20	-7.08	206.7
	Changes in pollution			
Total Sales (% point)	0.27	0.42	2.81	4.24
Emissions per vehicle %	0.04	0.62	0.41	5.45
	Changes in Profits			
Gamer/non gamer %	20.41	37.94	-9.86	-15.83

Column (1) and (3) give changes in market equilibrium when one firm games. We present the average of letting each firm game (we run a separate simulation per firm). Changes in profits are the average for the firms that game (not for the industry). Column (2) and (4) give changes in market equilibrium when all but one firm games. We run a separate simulations for each firm not gaming along when all others game. Changes in profits are the average for the firms that do not game along.

firm games, while all other firms remain honest. In the right panel, labeled “not gaming along,” all but one firm games. We simulate these asymmetric cases separately for each firm and present the average outcomes over these cases. In all scenarios we find results that are qualitatively similar to our industry-wide symmetric gaming case in that gaming hurts buyers when there is no policy, while buyers benefit from gaming under the policy. Buyer surplus and pollution effects are much smaller when only one of the carmakers games. Effects are almost equal to symmetric gaming when all but one firm games.

Table 3 also reports how profit changes with gaming. In the left panel we give the gains in profits for the single firm that games (the gamer), in the right panel we present the losses in profits for the single firm that does not game along (the non-gamer). The average gain from gaming unilaterally is large: a 20% increase in profits, even in the absence of policy. But the profitability from gaming rises still further, to 38%, with policy. This shows that firms have a very strong incentive to unilaterally deviate from honesty, and this incentive rises with policy. The profit losses from not “gaming along” are also large. Being the sole honest firm leads to a 10% decrease in profits without policy and a 16% decrease with policy. Thus, if competitors start to game, it is very costly not to follow along, and this incentive is enhanced by the policy.

We conclude that the policy thus clearly increases the incentive to game in a multi-firm context: the unilateral profits from deviating from honesty increase and the costs from not deviating when others do increases.²⁸ This is related to an existing literature that asks whether competition induces

²⁸Note that this setting is not necessarily a prisoner’s dilemma. The total industry profit from the price equilibrium with gaming relative to the price equilibrium with honesty will depend on the total legal and regulatory costs of gaming

gaming (e.g., [Becker and Milbourn 2011](#), [Bennett et al. 2013](#)), but we add to that the suggestion that corrective policy amplifies these competitive effects.

5 Conclusion

This paper does three things. First, it documents a remarkable and growing divergence between official and on-road fuel consumption rates, which we call the performance gap, in a reduced-form empirical analysis using novel data. This performance gap is coincident with the introduction of strict standards based on the official ratings, and we interpret this pattern as evidence of Goodhart’s Law, whereby policy-induced stakes lead to the strategic manipulation of a measure. Combined with the still-growing Volkswagen scandal regarding local air pollution, the performance gap in carbon emissions points to a veritable crisis in the administration of environmental regulations for automobiles. Moreover, combined with other findings that call into question the efficacy of other energy efficiency policies ([Davis, Fuchs, and Gertler 2014](#), [Fowlie, Greenstone, and Wolfram Forthcoming](#), [Levinson 2016](#)), we believe our findings point to a serious challenge for the future of climate negotiations, which are now built upon bottom-up pledges based on unreliable ex ante estimates of program conservation.

Second, we develop a theoretical framework in order to ask how such policy-induced gaming of energy-efficiency ratings will impact buyers. The model points out how the existence of an externality-correcting policy fundamentally alters the incidence of gaming on buyers. Absent policy, gaming hurts consumers through both the price effect and choice distortion channels. But, when gaming is done to avoid a regulation, consistent with Goodhart’s Law, then buyers can benefit via pass through of cost savings. We show that beneficial price effects can dominate so that private buyer surplus rises with gaming.

Third, we conduct a structural analysis and calibrated welfare simulation in order to quantify the price effects and choice distortions identified by theory. For a wide range of parameter values, we find that price effects do indeed dominate so that the theoretical result that consumers might benefit from gaming is not just a curious possibility, but in fact appears to be a likely probability.

Our analysis emphasizes private buyer surplus, abstracting from the externality. Assuming that the policies are optimally designed and the externality benefits accrue to buyers, the buyers would be better off when the seller complies honestly, once the externality is taken into account. But, in our empirical context, neither condition is likely to hold. Analysis of the fuel economy standards in the US typically finds that the standards do not pass a cost-benefit test unless buyers, contrary to recent empirical evidence, have large biases in their valuation of fuel economy. Moreover, the climate benefits of reduced emissions largely accrue to future generations and to people other than European car buyers.

We anticipate that the same concerns we highlight here are present in a variety of regulatory settings, and we expect that both our theoretical and empirical approach can be applied elsewhere.

relative to the total costs of compliance. The evidence presented here shows that honesty becomes harder to support when the policy is in place.

Critical to our setup is that the measure which can be gamed is used by the regulator *and* is instrumental to consumer choice. For example, in education, test scores of current students might be gamed to satisfy policy mandates that affect resource transfers. The test scores also affect the demand for admission from future waves of students. Consumers (students) can benefit from gaming when gaming frees up resources at a school that would otherwise be expended to satisfy policy and obtain transfers. Or, in finance, capital requirements can be used to minimize an individual bank's role in creating systemic risk. Customers can privately benefit if a bank is able to rebalance its portfolio more profitably after gaming the regulated measures.

References

- Allcott, Hunt. 2013. "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." American Economic Journal: Economic Policy 5 (3):30–66. URL <http://www.aeaweb.org/articles.php?doi=10.1257/pol.5.3.30>.
- Allcott, Hunt and Christopher Knittel. 2017. "Are Consumers Poorly Informed about Fuel Economy? Evidence from Two Experiments." NBER Working Paper 23076.
- Allcott, Hunt and Dmitry Taubinsky. 2015. "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market." American Economic Review 105 (8):2501–2538.
- Allcott, Hunt and Nathan Wozny. 2014. "Gasoline Prices, Fuel Economy, and the Energy Paradox." Review of Economics and Statistics 96 (5):779–795.
- Anderson, Soren T. and James M. Sallee. 2011. "Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards." American Economic Review 101 (4):1375–1409.
- . 2016. "Designing Policies to Make Cars Greener: A Review of the Literature." Annual Review of Resource Economics 8:157–80.
- Austin, David and Terry Dinan. 2005. "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes." Journal of Environmental Economics and Management 50 (3):562–582.
- Becker, Bo and Todd Milbourn. 2011. "How Did Increased Competition Affect Credit Ratings?" Journal of Financial Economics 101:493–514.
- Ben-David, Itzhak. 2011. "Financial Constraints and Inflated Home Prices during the Real Estate Boom." American Economic Journal: Applied Economics 3:55–87.
- Bennett, Victor Manuel, Lamar Pierce, Jason A. Snyder, and Michael W. Toffel. 2013. "Customer-Driven Misconduct: How Competition Corrupts Business Practices." Management Science 59 (8):1725–1742.
- Bento, Antonio M., Kenneth Gillingham, and Kevin Roth. 2017. "The Effect of Fuel Economy Standards on Vehicle Weight Dispersion and Accident Fatalities." Working Paper: Yale University.

- Bento, Antonio M., Lawrence H. Goulder, Mark R. Jacobsen, and Roger H. von Haefen. 2009. “Distributional and Efficiency Impacts of Increased US Gasoline Taxes.” American Economic Review 99 (3):667–699.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” Econometrica 64 (4):841–890.
- Booz & Company. 2012. “Company Cars and Fring Benefit Tax: Understanding the Impacts on Strategic Transport Targets.” NZ Transport Agency research report 474.
- Buchanan, James M. 1969. “External Diseconomies, Corrective Taxes, and Market Structure.” American Economic Review 59 (1):174–177.
- Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer. 2013. “Are Consumers Myopic? Evidence from New and Used Car Purchases.” American Economic Review 103 (1):220–256.
- Chetty, Raj, John N. Friedman, and Johan E. Rockoff. 2014. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” American Economic Review 104 (9):2593–2632.
- Davis, Lucas W., Alan Fuchs, and Paul Gertler. 2014. “Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico.” American Economic Journal: Economic Policy 6 (4):207–238.
- Dee, Thomas S., Will Dobbie, Brian A. Jacob, and Johan E. Rockoff. 2016. “The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations.” NBER Working Paper 22165.
- Drugov, Mikhail and Marta Troya-Martinez. 2015. “Vague Lies and Lax Standards of Proof: On the Law and Economics of Advice.” SSRN Working Paper.
- Dubois, Pierre, Rachel Griffith, and Martin O’Connell. 2018. “The Effects of Banning Advertising in Junk Food Markets.” The Review of Economic Studies 1 (1):396–436. URL <https://dl.dropboxusercontent.com/u/21832052/site/dgo.pdf>.
- Duflo, Esther, Michael Greenstone, and Nicholas Ryan. 2013. “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India.” Quarterly Journal of Economics 128 (4):1499–1545.
- Durrmeyer, Isis and Mario Samano. Forthcoming. “To Rebate or not to Rebate: Fuel Economy Standards vs. Feebates.” Economic Journal URL http://docs.wixstatic.com/ugd/23918f_a6dbf5dcc28046cabd57be5b8528b91e.pdf.
- Fabra, Natalia and Mar Reguant. 2014. “Pass-Through of Emissions Costs in Electricity Markets.” American Economic Review 104 (9):2872–2899.
- Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. Forthcoming. “Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program.” Quarterly Journal of Economics .
- Frankel, Alex and Navin Kartik. Forthcoming. “Muddled Information.” Journal of Political Economy .

- Gabaix, Xavier and David Laibson. 2006. “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets.” Quarterly Journal of Economics 121 (2):505–540.
- Gerlagh, Reyer, Inge van den Bijgaart, Hans Nijland, and Thomas Michielsen. Forthcoming. “Fiscal policy and CO2 emissions of new passenger cars in the EU.” Environmental and Resource Economics .
- Gillingham, Kenneth. 2013. “The Economics of Fuel Economy Standards versus Feebates.” Working Paper: National Energy Policy Institute.
- Goldberg, Pinelopi Koujianou. 1998. “The Effects of the Corporate Average Fuel Efficiency Standards in the U.S.” The Journal of Industrial Economics 46 (1):1–33.
- Goodhart, Charles. 1981. Problems of Monetary Management: The U.K. Experience. Springer.
- Goulder, Lawrence H. and Ian W. H. Parry. 2008. “Instrument Choice in Environmental Policy.” Review of Environmental Economics and Policy 2 (2):152–174.
- Gray, Wayne B. and Jay P. Shimshack. 2011. “The Effectiveness of Environmental Monitoring and Enforcement: A Review of the Empirical Evidence.” Review of Environmental Economics and Policy 5 (1):3–24.
- Grigolon, Laura, Mathias Reynaert, and Frank Verboven. Forthcoming. “Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market.” American Economic Journal: Economic Policy URL <http://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWVfbnxcmcFua3ZlcmJvfGd40jZiMzQ4YzUwYjhiNjdjOTk>.
- Heidhues, Paul, Botond Köszegi, and Takeshi Murooka. 2017. “Inferior Products and Profitable Deception.” Review of Economic Studies 84:323–356.
- Ito, Koichiro and James M. Sallee. Forthcoming. “The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards.” Review of Economics and Statistics .
- Jacobsen, Mark R. 2013. “Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior.” American Economic Journal: Applied Economics 5 (3).
- Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem. 2016. “Sufficient Statistics for Imperfect Externality-Correcting Policies.” Manuscript: University of California at Berkeley.
- Jacobsen, Mark R. and Arthur A. van Benthem. 2015. “Vehicle Scrappage and Gasoline Policy.” American Economic Review 105 (3):1312–1338.
- Jiang, Wei, Ashlyn Aiko Nelso, and Edward Vytlačil. 2014. “Liar’s Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency.” The Review of Economics and Statistics 96 (1):1–18.
- Kane, Thomas J. and Douglas O. Staiger. 2008. “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation.” NBER Working Paper 14607.
- Klier, Thomas and Joshua Linn. 2012. “New-vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard.” RAND Journal of Economics 43 (1):186–213.

- Kopczuk, Wojciech, Justin Marion, Erich Muehlegger, and Joel Slemrod. 2016. “Does Tax-Collection Invariance Hold? Evasion and the Pass-through of State Diesel Taxes.” American Economic Journal: Economic Policy 8 (2):1–36.
- Langer, Ashley and Shaun McRae. 2014. “Step on It: Evidence on the Variation in On-Road Fuel Economy.” Manuscript: University of Arizona.
- Leggett, Christopher G. 2002. “Environmental Valuation with Imperfect Information.” Environmental and Resource Economics 23:343–355.
- Levinson, Arik. 2016. “How Much Energy Do Building Energy Codes Save? Evidence from California Houses.” American Economic Review 106 (10):2867–2894.
- Lucas, Jr., Robert E. 1976. “Econometric Policy Evaluation: A Critique.” Carnegie-Rochester Conference Series on Public Policy 1:19–46.
- Oliva, Paulina. 2015. “Environmental Regulations and Corruption: Automobile Emissions in Mexico City.” Journal of Political Economy 123 (3):686–724.
- Parry, Ian W.H., David Evans, and Wallace E. Oates. 2014. “Are Energy Efficiency Standards Justified?” Journal of Environmental Economics and Management 67:104–125.
- Reynaert, Mathias. 2017. “Abatement Strategies and the Cost of Environmental Regulation: Emission Standards on the European Car Market.” Manuscript: Toulouse School of Economics.
- Reynaert, Mathias and Frank Verboven. 2014. “Improving the performance of random coefficients demand models: The role of optimal instruments.” Journal of Econometrics 179 (1):83–98. URL <https://ideas.repec.org/a/eee/econom/v179y2014i1p83-98.html>.
- Rhodes, Andrew and Chris M. Wilson. Forthcoming. “False Advertising.” RAND Journal of Economics .
- Roychowdhury, Sugata. 2006. “Earnings Management through Real Activities Manipulation.” Journal of Accounting and Economics 42:335–370.
- Sallee, James M. 2014. “Rational Inattention and Energy Efficiency.” Journal of Law and Economics 57 (3):781–820.
- Sallee, James M., Sarah E. West, and Wei Fan. 2016. “Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations.” Journal of Public Economics 135:61–73.
- Sanders, Nicholas J. and Ryan Sandler. 2017. “Technology and the Effectiveness of Regulatory Programs Over Time: Vehicle Emissions and Smog Checks with a Changing Fleet.” NBER Working Paper 23966.
- Sigman, Hilary. 2012. “Monitoring and Enforcement of Climate Policy.” In The Design and Implementation of U.S. Climate Policy, edited by Don Fullerton and Catherine Wolfram. University of Chicago Press, 213–225.
- Slemrod, Joel. 2008. “Does It Matter Who Writes the Check to the Government? The Economics of Tax Remittance.” National Tax Journal 61 (2):251–275.

- Small, Kenneth A. and Harvey S. Rosen. 1981. "Applied welfare economics of discrete choice models." Econometrica .
- Tanaka, Shinsuke. 2017. "Mind the Gap! Tax Incentives and Incentives for Manipulating Fuel Efficiency in the Automobile Industry." Working Paper: Tufts University.
- The International Council on Clean Transportation. 2014. "From Laboratory to Road: A 2014 Update of Official and "Real-World" Fuel Consumption and CO₂ Values for Passenger Vehicles in Europe." ICCT White Paper.
- . 2015. "From Laboratory to Road: A 2015 Update of Official and "Real-World" Fuel Consumption and CO₂ Values for Passenger Cars in Europe." ICCT White Paper.
- Walker, Reed W. 2013. "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce." Quarterly Journal of Economics 128 (4):1787–1835.
- Weyl, E. Glen and Michael Fabinger. 2013. "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition." Journal of Political Economy 121 (3):528–583.
- Whitefoot, Kate S., Meredith L. Fowlie, and Steven J. Skerlos. 2017. "Compliance by Design: Influence of Acceleration Trade-offs on CO₂ Emissions and Costs of Fuel Economy and Greenhouse Gas Regulations." Environmental Science & Technology 51 (18):10307–10315.
- Whitefoot, Katie and Steven Skerlos. 2012. "Design Incentives to Increase Vehicle Size Created from the U.S. Footprint-based Fuel Economy Standards." Energy Policy 41:402–411.
- Zitzewitz, Eric. 2012. "Forensic Economics." Journal of Economic Literature 50 (3):731–769.

A APPENDIX MATERIAL

A.1 Sample selection process:

Self-reported odometer readings are sometimes missing or clearly incorrect. We eliminate unreliable data through a data selection process detailed here.

Our data selection process first eliminates a number of drivers (accounts) that have unreliable information or too few observations. With the accounts that remain, we then consider several ways to account for mismeasurement in odometer readings for individual transactions. Specifically, we first limit the sample to gasoline- or diesel-fueled vehicles, which eliminates 6.7 million transactions. Second, we drop vehicles that use the wrong type of fuel for their engine in more than 1% of the visits, e.g., putting diesel fuel in a vehicle that is labeled as gasoline in our data. Inconsistencies might be in the data because drivers use their card for a different vehicle, or these observations might be mistakes in the assignment of vehicle type. This drops 7.5 million transactions. Third, we pose some minimum requirements on the driving patterns of the drivers that produce the transactions. We drop drivers that never report an increase of more than 150km in their odometer reading (2.5 million transactions).²⁹ We drop car models with fewer than 10 drivers, and drivers with fewer than 10 fuel station visits (1.3 million transactions). We drop drivers that did not report driving more than 5,000km in total or reported driving more than 500,000km in total (11.3 million). Having isolated a set of drivers (accounts) with ample data, within those accounts we drop individual transactions in two steps. First, we drop transactions where the odometer difference is lower than 100km or higher than 3000km (7.6 million). Second, we drop transactions that result in a fuel consumption that is outside 1.25 times the interquartile range of estimated fuel consumption for each car model in the data (5.1 million). This results in the final dataset of 24 million observations.

A.2 Empirical Bayes Correction of on-road ratings

We start by decomposing the total variance in the sample $Var(r_{nij}) = \sigma_r^2$ into three components: variance in performance of vehicles σ_j^2 , drivers σ_i^2 , and pump visits σ_n^2 . We estimate the variance between pump visits of the same driver as:

$$\sigma_n^2 = \frac{1}{N-I} \sum_n (r_{nij} - \bar{r}_{ij})^2,$$

in which \bar{r}_{ij} is the mean fuel consumption of driver i , N is the total number of observations and I is the total number of drivers. Next, we estimate the covariance between drivers of the same vehicles as:

$$\sigma_j^2 = cov(\bar{r}_{ij}, \bar{r}_{kj}).$$

²⁹Note that the range of a combustion engine is easily more than 800km.

The estimated covariance is obtained as a weighted average of covariances between randomly sorted pairs (i, k) of drivers of the same car. We weigh each pair of drivers (i, k) by the sum of their visits. Finally, we obtain σ_i^2 as the remaining variance: $\sigma_i^2 = \sigma_r^2 - \sigma_n^2 - \sigma_j^2$. The precision of the estimated gap for each driver is then defined as:

$$h_i = 1/(\sigma_i^2 + \sigma_n^2/n_i),$$

so that drivers with a high number of visits have a higher precision. We obtain precision weighted means per car as the weighted average of \bar{r}_{ij} with h_i as weights. Second we shrink these precision weighted means with an estimate of their reliability:

$$\psi_j = \sigma_j^2/(\sigma_j^2 + 1/\sum_i h_{ji}),$$

where the reliability is defined as the signal σ_j over the total variance. We use the per vehicle shrunk on-road estimates \hat{r}_j to construct an alternative estimate of the gap defined in (2) and to inform us about the distribution of the gap between vehicles.

Table A.3 describes the variation in the on-road fuel consumption (σ_d^2) across release years, as well as its decomposition across three components: variation across refueling transactions for the same driver (σ_n^2), variation across drivers of the same vehicle (σ_i^2) and variation across vehicles (σ_j^2). We decompose the variation separately for each release year and describe the mean and standard deviation across release years in the table.

More than 25% of the variance is attributable to within driver variance. This variance is due to driving conditions, stockpiling effects and errors in odometer reporting. We find that the variance across drivers of the same car σ_i^2 is 0.21. This is an economically large number; it means that the on-road fuel consumption is estimated to be 0.28 liter/100km higher at the third quartile than at the first quartile of drivers in the same car.³⁰ A policy that would shift a driver from the third quartile of the fuel consumption gap to the first quartile would decrease fuel consumption by 3%. These numbers are interesting from a policy perspective as they give an indication of the extent to which fuel consumption and emissions can be reduced by teaching and incentivizing drivers to drive a vehicle more efficiently.³¹ The remaining part of the variance σ_j^2 is the co-variance between drivers of the same car and can be seen as the information available to estimate the car specific component of on-road fuel consumption. We estimate this to be 1.35, which is more than 60% of the total variance. Table A.3 also shows that the variance components are relatively stable over time; each component has a low standard deviation across release years. There is variation in the size of the fuel consumption gap between cars and between drivers, but this variation is stable over time. Given this variance decomposition we turn next to the estimates of the distribution

³⁰If we assume that conditional on car j , r has a normal distribution, the interquartile distance is $1.349*\sigma_i^2$.

³¹Significant variation across drivers of identical cars is consistent with results reported in Langer and McRae (2014), who analyze extremely detailed driving data from a few dozen drivers of an identical car, the Honda Accord. In contrast, our data come from a large sample and cover many models.

of r_j and d_j for each release year. Table A.4 reports the unweighted mean estimate of r_j and \hat{r}_j , obtained with the empirical Bayes correction. The mean value of both r_j and \hat{r}_j are decreasing over the release years. In all years the corrected means are lower than the raw means, because on average vehicles with high r_j have less precise underlying data, but overall shrinkage and precision weighting has small effects. The resulting gap \hat{d}_j is estimated to be an imprecise 10% up until 2006. From 2007 onwards we see a significant increase in the performance gap, consistent with the previous estimates.

A.3 Proofs

Proposition 1. *In the absence of policy ($\lambda = 0$), buyer surplus falls with the level of gaming. Specifically:*

$$\frac{dBS}{dg} \approx \underbrace{-\rho(1-\alpha)\beta D}_{\text{price effect}} + \underbrace{(1-\alpha)^2\beta^2 D'g}_{\text{choice distortion}} \leq 0.$$

We derive this by starting with the definition of buyer surplus as the integral under the inverse demand curve. Because the true attribute of the good x is unaffected by gaming, we can analyze buyer surplus using the true demand curve. The standard portion of buyer surplus is the integral from the final price, denoted p^* up to infinity. Denote by \tilde{p} the upfront purchase price that would induce a sophisticated consumer to purchase the amount of the good that is in fact purchased at price p^* by the consumer with perception $(1-\alpha)$. The choice distortion can be written as the difference between the revenue generated between p^* and \tilde{p} and the consumer value generated between those points.

$$BS = \underbrace{\int_{p^*}^{\infty} D(z + \beta x) dz}_{\text{BS of correct quantity}} + \underbrace{\int_{\tilde{p}}^{p^*} D(z + \beta x) dz}_{\text{Value of excess quantity}} - \underbrace{\int_{\tilde{p}}^{p^*} D(\tilde{p} + \beta x) dz}_{\text{Cost of excess quantity}}. \quad (\text{A.1})$$

Differentiation of equation A.5, in which p^* and \tilde{p} are endogenously determined by g , yields the result. Note that the inside of the third integral is a constant with respect to the variable of integration, so it can be pulled out of the integral, leaving only the constant 1 inside. Specifically:

$$\begin{aligned} \frac{dBS}{dg} = & -D(p^* + \beta x) \frac{dp^*}{dg} + \left\{ D(p^* + \beta x) \frac{dp^*}{dg} - D(\tilde{p} + \beta x) \frac{d\tilde{p}}{dg} \right\} \\ & - \left\{ (p^* - \tilde{p}) D'(\tilde{p} + \beta x) \frac{d\tilde{p}}{dg} + D(\tilde{p} + \beta x) \left(\frac{dp^*}{dg} - \frac{d\tilde{p}}{dg} \right) \right\}. \end{aligned} \quad (\text{A.2})$$

Using the pass through coefficient ρ , a change in g scales to a change in tax by $(1-\alpha)\beta$, so $dp^*/dg = \rho(1-\alpha)\beta$. This simplifies the first term to yield the result.

For the second term in A.2, note that $\tilde{p} = p^* - (1-\alpha)\beta g$ by definition. Then, substitute a first-

order Taylor approximation to write demand at \tilde{p} as a function of demand at p^* and $D' \times (\tilde{p} - p^*)$:

$$\begin{aligned}
& D(p^* + \beta x) \frac{dp^*}{dg} - D(\tilde{p} + \beta x) \frac{d\tilde{p}}{dg} \\
&= D(p^* + \beta x) \frac{dp^*}{dg} - D(p^* + \beta x - (1 - \alpha)\beta g) \frac{d\tilde{p}}{dg} \\
&\approx \{D(p^* + \beta x - (1 - \alpha)\beta g) + D'(p^* + \beta x - (1 - \alpha)\beta g)(1 - \alpha)\beta\} \frac{dp^*}{dg} - D(p^* + \beta x - (1 - \alpha)\beta g) \frac{d\tilde{p}}{dg} \\
&= D(p^* + \beta x - (1 - \alpha)\beta g) \left(\frac{dp^*}{dg} - \frac{d\tilde{p}}{dg} \right) + D'(p^* + \beta x - (1 - \alpha)\beta g)(1 - \alpha)\beta \frac{dp^*}{dg} \tag{A.3}
\end{aligned}$$

Now consider the third term in [A.2](#). After substituting $\tilde{p} = p^* - (1 - \alpha)\beta g$, we see that the term that multiplies the difference in derivatives will cancel in the third term of [A.2](#) and the second-term, defined using [A.3](#). This means that [A.2](#) can be written:

$$\begin{aligned}
\frac{dBS}{dg} &= -D(p^* + \beta x)\rho(1 - \alpha)\beta + D'(p^* + \beta x - (1 - \alpha)\beta g)(1 - \alpha)\beta \frac{dp^*}{dg} \\
&\quad - \left\{ (p^* - \tilde{p})D'(\tilde{p} + \beta x) \frac{d\tilde{p}}{dg} \right\}. \tag{A.4}
\end{aligned}$$

Substitute $p^* - \tilde{p} = (1 - \alpha)\beta g$ and $d\tilde{p}/dg = (\rho - 1)(1 - \alpha)\beta$, which follows from differentiating the definition of \tilde{p} and using the pass through result for p^* . Then, simplification yields the final result:

$$\begin{aligned}
\frac{dBS}{dg} &= -D(p^* + \beta x)\rho(1 - \alpha)\beta + D'(p^* + \beta x - (1 - \alpha)\beta g)(1 - \alpha)\beta\rho(1 - \alpha)\beta \\
&\quad - \{(1 - \alpha)\beta g D'(\tilde{p} + \beta x)(\rho - 1)(1 - \alpha)\beta\} \\
&= -D(p^* + \beta x)\rho(1 - \alpha)\beta + D'(p^* + \beta x - (1 - \alpha)\beta g)(1 - \alpha)^2\beta^2\rho \\
&\quad - (1 - \alpha)^2\beta^2 g D'(\tilde{p} + \beta x)(\rho - 1) \\
&= -D(p^* + \beta x)\rho(1 - \alpha)\beta + (1 - \alpha)^2\beta^2 g D'(\tilde{p} + \beta x).
\end{aligned}$$

Note that, with the local linear demand assumption, the derivative of D evaluated at either \tilde{p} or p^* is the same. ■

Proposition 2. *In the presence of a binding standard ($\lambda > 0$), a change in gaming affects buyer surplus as follows:*

$$\frac{dBS}{dg} \approx \underbrace{(-\rho(c' - \alpha\beta) - \beta)D}_{\text{price effect}} + \underbrace{(1 - \alpha)^2\beta^2 D'g}_{\text{choice distortion}}.$$

Proposition 2 is derived in a similar way to Proposition 1, but we define surplus using integrals over the demand function starting with full prices f . (This same could have been done in the prior proof, yielding the same result.) Recall that $f = p + \beta x$ and $\tilde{f} = p + \beta x - (1 - \alpha)\beta g$.

$$BS = \underbrace{\int_{f^*}^{\infty} D(z)dz}_{\text{BS of correct quantity}} + \underbrace{\int_{\tilde{f}}^{f^*} D(z)dz}_{\text{Value of excess quantity}} - \underbrace{\int_{\tilde{f}}^{f^*} D(\tilde{f})dz}_{\text{Cost of excess quantity}}. \quad (\text{A.5})$$

Differentiating, using the same Taylor approximation as above to simplify, yields:

$$\frac{dBS}{dg} \approx -D(f^*)\frac{df^*}{dg} + D'(f^*)(f^* - \tilde{f}) \left(\frac{df^*}{dg} - \frac{d\tilde{f}}{dg} \right).$$

Substitute the pass through result described in the text: $df^*/dg = \rho(c'(x) + \alpha\beta) + \beta$. Substitute the definition of \tilde{f} , which shows that $f^* - \tilde{f} = (1 - \alpha)\beta g$. And substitute $\frac{df^*}{dg} - \frac{d\tilde{f}}{dg} = (1 - \alpha)\beta$:

$$\begin{aligned} \frac{dBS}{dg} &\approx -D(f^*)(\rho(c'(x) + \alpha\beta) + \beta) + D'(f^*) \{(1 - \alpha)\beta g\} \{(1 - \alpha)\beta\} \\ &= -D(f^*)(\rho(c'(x) + \alpha\beta) + \beta) + D'(f^*)(1 - \alpha)^2\beta^2 g. \end{aligned} \quad (\text{A.6})$$

This yields the result. ■

Proposition 3. *In the presence of a fuel tax but not a binding standard ($\lambda = 0, \tau > 0$), buyer surplus falls with the level of gaming. Specifically:*

$$\frac{dBS}{dg} \approx \underbrace{-\rho(1 - \alpha)(\beta + \tau)D}_{\text{price effect}} + \underbrace{(1 - \alpha)^2(\beta + \tau)^2 D'g}_{\text{choice distortion}} \leq 0.$$

Proposition 3 follows directly from Proposition 1 once it is recognized that this is the exact same problem with β replaced by $\beta + \tau$ at all points. Thus, the exact same proof as Proposition 1 holds, yielding the result. ■

A.4 Estimation Details

We estimate the parameters θ using the following GMM objective:

$$\min_{\theta} \xi_j Z' \omega Z \xi_j$$

in which Z is a matrix of instruments and ω is a weighting matrix.

We use a panel containing sales, prices and characteristics for all new vehicle sales in seven European countries between 1998 and 2007.³² We only use data from before 2008 to estimate demand, so that our estimates come from a period in which the performance gap was stable. We thus assume variation in x_j is informative about actual fuel cost differences. The vector Δ_j

³²The seven countries are: Belgium, France, Germany, Great Britain, Italy, Netherlands and Spain.

contains information on horsepower, weight, footprint (a measure of vehicle size), height and a dummy specifying if the car is of a foreign brand (e.g., Fiat in France). Additionally, we include fuel type by market dummies, dummies for the number of months a vehicle was on sale in a country-year, country fixed effects, a linear time trend, body type fixed effects, vehicle class fixed effects and brand fixed effects. We divide prices by income per capita in each country-year, so that price sensitivity varies with income in the market. We need instruments for price and for the standard deviations of the random coefficients. We instrument for prices using both cost shifters and sums of characteristics instruments, which follows [Berry, Levinsohn, and Pakes \(1995\)](#). The sums of characteristics instruments are the sum of fuel costs, horsepower, weight, footprint and height across all other products in the market and across all other products within the same firm in the market. We also include the number of competing products and the number of products in the same firm. The cost shifters are the log of labor costs in the country of production and a dummy specifying if the vehicle is sold in the country of production. For the standard deviations, we use approximately optimal instruments that are constructed using a two-step procedure as described in [Reynaert and Verboven \(2014\)](#). We estimate considerable heterogeneity in the taste for horsepower, weight and footprint. Vehicles perceived as foreign are less attractive for consumers. Cars with four doors are preferred over cars with two doors.

A.5 Additional Tables

Table A.1: Summary statistics

	Mean	St. Dev.
Car Characteristics		
Fuel Consumption (L/100km)	6.65	1.73
Vehicle Weight in kg	1,354	245
Diesel Engines	0.46	0.50
Drivers per car	107	219
Driver Characteristics		
Pump visits	134	80
Total liters purchased	6,015	3,666
Total distance (km)	111,726	53,942
Pump Visit Characteristics		
Liters per visit	45.3	10.8
Odometer increase per visit	671	192

The table gives summary statistics for the 2,696 vehicles, 266,616 drivers and 23,989,576 pump visits in the sample.

Table A.2: Summary statistics: Netherlands and Travelcard

	Mean	St. Dev.	Mean	St. Dev.
	TravelCard		Netherlands	
Price (euro)	31,672	13,367	40,767	29,676
Fuel Consumption (L/100km)	6.74	1.60	7.89	2.46
Vehicle Weight in kg	1,344	230	1,409	308
Diesel Engines	0.45	0.50	0.36	0.48

Summary statistics for the TravelCard sample and the full dutch market between 1998 and 2011.

Table A.3: Variance decomposition

	σ_d^2	σ_i^2	σ_j^2	σ_n^2
Mean	2.11	0.21	1.35	0.56
Standard deviation	0.57	0.03	0.45	0.11
Variance decomposition (%)	100	10.36	62.34	27.30
Standard deviation		2.54	6.76	4.40

σ_d^2 is the total variance in r_{nij} , σ_i^2 is the variation attributable to differences across individuals driving the same vehicle, σ_j^2 is the covariance between drivers in the same vehicle, σ_n^2 is the variation across refueling visits of the same driver in the same vehicle. The variance decomposition is performed separately for each release year, and the mean and standard deviation across years are reported in table.

Table A.4: Estimated vehicle gaps by release year, with and without empirical Bayes correction

Release	Unweighted r_j	\hat{r}_j	Shrinkage	\hat{d}_j
1998	7.91 (1.41)	7.87 (1.39)	0.99 (0.00)	7.85 (7.82)
1999	7.90 (1.49)	7.88 (1.46)	0.99 (0.01)	8.70 (6.90)
2000	7.96 (1.44)	7.94 (1.43)	0.99 (0.01)	8.60 (7.36)
2001	7.73 (1.44)	7.72 (1.41)	0.99 (0.00)	11.15 (7.12)
2002	7.63 (1.58)	7.60 (1.57)	0.99 (0.00)	11.02 (9.17)
2003	7.83 (1.61)	7.80 (1.56)	0.99 (0.00)	11.01 (9.18)
2004	8.31 (1.68)	8.29 (1.64)	0.99 (0.00)	9.52 (9.58)
2005	7.87 (1.71)	7.82 (1.59)	0.99 (0.01)	12.56 (9.14)
2006	8.10 (1.57)	8.06 (1.53)	0.99 (0.00)	12.55 (8.17)
2007	7.71 (1.37)	7.68 (1.34)	0.99 (0.00)	17.66 (9.67)
2008	7.46 (1.36)	7.45 (1.34)	0.99 (0.00)	17.57 (9.00)
2009	7.24 (1.30)	7.21 (1.28)	0.99 (0.00)	22.34 (10.07)
2010	7.15 (1.33)	7.13 (1.32)	0.99 (0.00)	26.38 (10.68)
2011	6.92 (1.24)	6.90 (1.21)	0.99 (0.01)	29.74 (10.29)
2012	6.64 (1.08)	6.62 (1.06)	0.99 (0.01)	37.46 (11.05)
2013	6.23 (0.92)	6.23 (0.89)	0.99 (0.01)	44.23 (11.34)
2014	6.18 (1.20)	6.15 (1.13)	0.98 (0.02)	52.90 (12.04)

Table reports mean and standard deviations for the distribution of estimated on-road consumption r_j , Bayes corrected on-road consumption \hat{r}_j , shrinkage factor and Bayes corrected efficiency gap \hat{d}_j by release year.

Table A.5: Estimation Results

	Mean Taste		St. Dev.	
	Coeff.	St. Error	Coeff.	St. Error
Price/Inc.	-6.51	(0.45)		
Fuel Cost	-0.53	(0.01)	0.05	(0.03)
Horsepower	1.48	(0.21)	1.78	(0.10)
Weight	0.22	(0.21)	4.32	(0.16)
Footprint	0.88	(0.05)	0.58	(0.04)
Foreign	-0.92	(0.03)	0.02	(0.16)
Height	0.02	(0.02)		
Doors	0.50	(0.11)		

The table shows estimated taste parameters from a random coefficient logit estimation on the the car market for seven EU countries using data from 1998 to 2007. Taste distributions are assumed to be normal, and mean and standard deviations are estimated for selected characteristics. Additional controls are fuel type by market dummies, months for sale if less than 12, country fixed effects, linear time trend, body type fixed effects, vehicle class fixed effects and brand fixed effects. Model is estimated using a two-step GMM using approximate optimal instruments with sum of characteristics and cost shifter instruments for prices.

Table A.6: Consumer surplus changes from gaming under alternative policies

	Policy Target		
	3%	5%	10%
	Attribute-Based Standard		
Total	33.19	208.27	608.74
Choice Distortion	-17.77	-16.72	-13.47
Price Effect	50.96	224.99	622.22
	Flat Standard		
Total	3.92	107.93	398.69
Choice Distortion	-17.58	-16.93	-15.01
Price Effect	21.50	124.86	413.70
	Fuel Tax		
Total	-24.59	-24.94	-25.80
Choice Distortion	-17.42	-17.66	-18.23
Price Effect	-7.16	-7.29	-7.57

This Table presents consumer surplus changes when there is industry wide 5% gaming and policies with varying stringency: in Column (1) the target is a 3% reduction in emissions, Column (2) a 5% reduction (as in the previous Tables) and Column (3) a 10% reduction. The first panel shows the consumer surplus changes when firms game and there is an attribute-based standard, the second panel shows the effect of gaming when there is a flat standard, the third panel shows the effect of gaming when the policy is a fuel tax rather than a standard.