# The Implications of Heterogeneity for the Regulation of Energy-Consuming Durable Goods

Mark R. Jacobsen,[1,5] Christopher R. Knittel,[2,5] James M. Sallee,[3,5] Arthur A. van Benthem[4,5]*

March 22, 2015

**Abstract**

Many of the most important policies that aim to reduce greenhouse gas emissions and other environmental externalities do so by regulating the energy efficiency ratings of energy-consuming durable goods. However, each product's lifetime externality depends not only on these ratings, but also on its lifetime utilization. As a result, conventional energy efficiency policies are inefficient when products differ significantly in their average longevity. We develop a theoretical model that characterizes this inefficiency using sufficient statistics that require minimal market data. We then explore the quantitative importance of this phenomenon for the case of automobiles using data on lifetime vehicle mileage from a large sample of automobiles. We document substantial heterogeneity in the longevity of different types of cars, and our model translates this heterogeneity into welfare implications. We estimate that fuel economy standards that regulate fuel economy but ignore longevity are able to recover only about one-quarter to one-third of the welfare gains achievable by a policy that also takes longevity into account.

Keywords: Corrective taxation, externalities, heterogeneity, sufficient statistics
JEL: H23, Q58, L51

*We thank Don Fullerton, Jeff Grogger and seminar participants at the AEA meetings, UC Berkeley and the University of Chicago for helpful comments and suggestions.

[1]Department of Economics, University of California San Diego. [2]William Barton Rogers Professor of Energy Economics, Sloan School of Management, Director, Center for Environmental and Energy Policy Research, Massachusetts Institute of Technology. [3]Harris School of Public Policy, University of Chicago. [4]The Wharton School, University of Pennsylvania. [5]National Bureau of Economic Research. 

# 1 Introduction

The consumption of energy is nearly always achieved through the operation of some durable good. Motor vehicles combust gasoline; appliances use electricity; furnaces burn natural gas; and so on. To correct the market failures caused by the pollution that attends energy consumption, economists typically advocate the pricing of emissions. As argued by Pigou (1932), if externalities can be taxed directly, then market efficiency can be fully restored. Such policy prescriptions are indifferent to the durables that act as an intermediary between fuel inputs and emissions outputs.

Policies that directly target emissions are, however, relatively rare. Instead, a proliferation of policies focus on these durable intermediaries, often through the regulation of their energy efficiency. Examples include fuel economy regulations, appliance efficiency mandates or standards, and building codes. It is well established that energy efficiency policies suffer from inefficiencies both because they fail to incentivize abatement on the intensive margin and because, even if they get relative prices of goods right within a market, they typically fail to set the average price level correctly.[1] The aim of this paper is to establish another inefficiency of such policies, one that stems from heterogeneity in how long durables last.

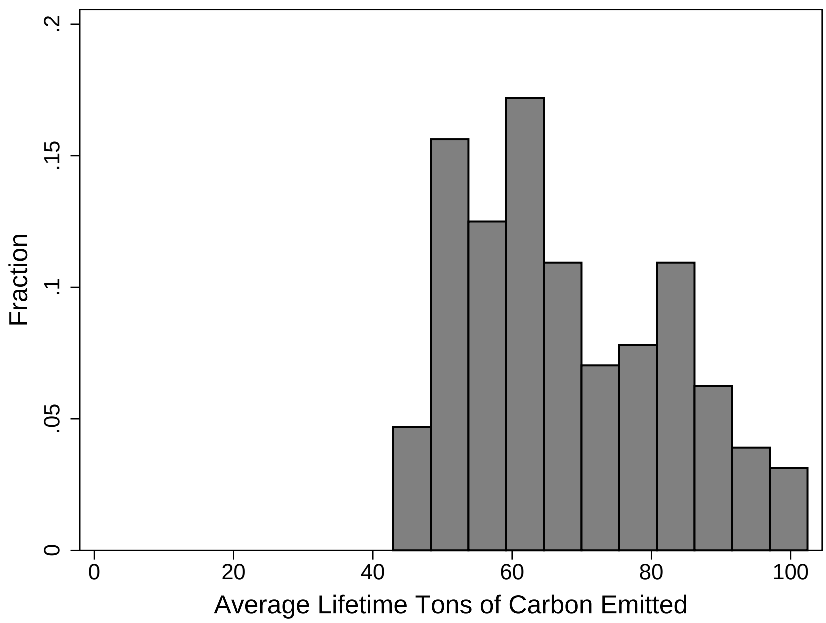
To see the logic of our inquiry, consider a consumer who is to buy one vehicle and will drive that vehicle a fixed number of miles per year. Her only choice is which of several vehicles to buy. The vehicles have different expected longevities, due to differences in their quality. Differences in longevity imply that two vehicles with the same fuel economy will nevertheless have different lifetime emissions. A gasoline tax (or a carbon tax) would raise the lifetime expected fuel cost of each model according to both its fuel economy and its expected longevity. In contrast, policies that regulate the durables themselves, rather than emissions (or fuel inputs), are almost always unable to take such heterogeneity in durability into account. Rather, policy treatment generally depends exclusively on the durable good's energy efficiency rating (e.g., fuel economy). This limits the ability of durable goods regulations to induce consumers to purchase the socially optimal products, which creates the inefficiency that is the subject of our investigation.

Broadly, heterogeneity in the utilization of a durable can come from *ex ante* differences in the product's quality, *ex ante* differences in the intensity of use across consumers, or *ex post* realizations of random product failure. Random failure that is independent of choice will not influence the relative efficiency of one policy versus another, and so we focus on the other two sources of heterogeneity. The heart of our analysis is focused on the first source: *ex ante* differences in product quality that predict longevity.

To illustrate the importance of this source of heterogeneity for an example durable, automobiles,

[1] For example, a gasoline tax would raise the price of driving and thus reduce automobile usage, and it would raise the cost of ownership for all cars, thereby shrinking the car market overall. In contrast, fuel economy regulations lower the cost of driving, and they implicitly subsidize efficient cars while taxing inefficient cars, which fails to optimally shrink the market. See Anderson, Parry, Sallee, and Fischer (2011) and Sallee (2011), respectively, for discussions of the efficiency of fuel economy regulations and taxes. See Borenstein (Forthcoming) for a recent treatment of the economics of the rebound effect, which relates to the intensive margin issue. See Holland, Hughes, and Knittel (2009) for an exploration of how performance standards create inefficiencies due to their average price effects.

**Figure 1:** Distribution of Lifetime Carbon Emissions for Vehicles with EPA Average Rating of 23 Miles per Gallon




An observation is the average lifetime miles driven of a particular type of vehicle, across many individual units, divided by fuel economy rating multiplied by the tons of carbon per gallon of gasoline. The sample is restricted to models for which we observe at least 200 vehicle retirements from model years 1988 to 1992, and to cars with a fuel economy rating between 23 and 24 MPG. The data are described in detail below.

we plot the average lifetime carbon emissions for different types of cars that have the same fuel economy rating in Figure 1; i.e., we illustrate the variance in lifetime pollution for products with identical energy efficiency ratings. The figure uses data, which we describe in detail below, on the odometer readings of vehicles shortly before they are scrapped, which we convert into tons of carbon using the vehicle's fuel economy rating and the average carbon content of gasoline. Each data point is the average lifetime carbon emissions across a number of individual vehicles of the same type (e.g., a 2012 Toyota Camry). The graph suggests a wide dispersion in lifetime emissions among vehicles with a common 23 miles per gallon rating (the median in our data); the standard deviation is 20% of the median. At $39 per ton, which is the current federal guideline for the social cost of carbon, the standard deviation in damages across cars with the same fuel economy rating is over $600. Thus, a durable goods policy, like Corporate Average Fuel Economy (CAFE) standards, that gives all cars with the same fuel economy rating the same implicit regulatory shadow price is necessarily imprecise; it places the same implicit tax or subsidy on products that in fact have substantially different lifetime externalities.

We are not the first to consider how heterogeneity affects the welfare properties of policies that regulate durable goods, as compared to efficient policies that target emissions. Our paper is most related to Fullerton and West (2010).[2] Using data on vehicle ownership, miles driven, and emissions, Fullerton and West (2010) simulate the welfare improvement from the optimal set of Pigouvian taxes as well as a simple gas tax. While their focus is not on the heterogeneity in externalities due to variation in the durability of products, their analysis does account for other important sources of heterogeneity such as vehicle-specific differences in emissions rates per mile of local air pollutants. A major advantage of the simulation exercise in Fullerton and West (2010), and similar papers that either simulate or structurally estimate market equilibria, is that they are able to study the welfare implications of a rich set of policies.

Rather than emphasizing a particular model of the market, we take a sufficient statistics approach that allows us to make welfare statements with a minimal amount of market data and only a couple of key parameters. Our focus is on developing an intuitive theoretical structure that can clarify how heterogeneity in utilization affects the welfare properties of a variety of important policies. These policies can be mapped into a tax schedule, which can be any linear or non-linear function of product attributes such as energy efficiency ratings. We derive a sufficient statistic for the welfare consequences of such second-best policies that tax products based only on their energy efficiency. Therefore, our paper is also related to papers such as Chetty (2009) that bridge the gap between structural estimation and simulation and reduced form results.

Our paper does three things. First, we develop a representative consumer model—within which heterogeneity is driven entirely by product quality—and derive sufficient statistics for the dead-

[2] A number of other papers, including Fullerton and West (2002), Feng, Fullerton, and Gan (2013), Holland, Mansur, Muller, and Yates (2014) and Knittel and Sandler (2013), also explore the implications of heterogeneity in local air pollution from automobiles for policy design. Another strain of literature explores heterogeneity in emissions reductions in electricity generation from renewables or demand side management, including Cullen (2013) and Callaway, Fowlie, and McCormick (2015). Our paper differs from all of these in using the sufficient statistics approach.

weight loss of using regulations based solely on energy efficiency ratings in lieu of policies, such as an emissions tax, that account for differences in product longevity. To highlight the role of heterogeneity, we assume that intensive margin utilization is exogenous. In this case, the first-best allocation can be achieved by a set of new product taxes that raises the price of each product according to its *ex ante* expected lifetime external damages, which depend on both energy efficiency ratings and average longevity. In contrast, a regulatory policy that mandates a market-wide average energy efficiency rating (like CAFE does for cars), is constrained to implement an implicit tax schedule that is a linear function of energy consumption ratings. Our framework illustrates that, under some plausible conditions, the fraction of the welfare gain achieved by the first-best policy—as compared to a baseline policy that sets a single tax rate on all products—that can be achieved by a linear tax on energy efficiency is equal to the $R^2$ from a regression of the lifetime emissions of products on their energy efficiency ratings. Intuitively, the first-best policy assigns each product a tax equal to its lifetime damages. The ability of the linear tax to mimic this differentiation is determined by the degree to which efficiency ratings predict lifetime emissions. This result links directly to graphs like Figure 1; when such figures show a greater spread in lifetime emissions for a given energy efficiency rating, the $R^2$ of a regression of lifetime emissions on energy efficiency ratings will be lower and the policy's deadweight loss will be greater. Under alternative assumptions, terms other than this $R^2$ will also influence welfare, but the $R^2$ remains an important, often dominant, factor.

Second, we demonstrate the quantitative importance of our theoretical results using the case of automobiles. We use data from the California Smog Check Program, paired with an industry source that identifies when a vehicle has been retired from the U.S. fleet, to estimate the average lifetime miles driven for many types of automobiles. We document evidence of very large differences in longevity across different makes and models. We use our data to calculate the welfare statistics derived in our theoretical model and conclude that, relative to a constant tax on all automobiles, policies that ignore heterogeneity in longevity (such as fuel-economy standards) recover only about one-quarter to one-third of the welfare gains that are achievable by the efficient policy.

Third, we investigate the robustness of our sufficient statistics in a more general model that accounts for heterogeneity across types of consumers, who may use durable goods differently. In this setting, the simple $R^2$ will no longer be a sufficient statistic. Our results demonstrate that a vast amount of information is required in order to formulate the welfare implications of second-best policies. In particular the remaining deadweight loss will depend on a matrix of interactions between cross-product price elasticities and marginal external damages.[3] We plan to develop simulations that allow us to investigate the accuracy of using the simple $R^2$ measure to approximate welfare in this more general setting.

[3] Our analysis in this section is related to Diamond (1973), which considers the second-best tax rates for a single consumption good that causes a different externality when consumed by different individuals. Our model generalizes the results of Diamond (1973) to the case of many interrelated goods (e.g., types of cars) that each cause externalities.

## 2 Model with a representative consumer

We begin with a representative consumer model. This allows us to isolate the welfare implications of heterogeneity in externalities caused by inherent differences across products, rather than heterogeneity caused by different consumers who buy identical products but use them differently. We make several assumptions that allow us to highlight the implications of this type of heterogeneity. First, we assume that intensive margin utilization of each product type is exogenous to the policy regime. For example, in the car market, we would assume that the average lifetime mileage of each type of car is fixed. This abstracts from differences in outcomes between a gas tax and fuel economy standard that are due to the intensive margin (rebound effect). The second key assumption is that the characteristics of the products are fixed. That is, this is a short-run perspective that does not include endogenous technology adoption or the response of other product attributes to policy.

With a representative consumer and exogenous intensive utilization and exogenous technology, each product will have a fixed lifetime social cost, and the only thing that determines welfare is the consumer's portfolio choice of goods. Thus, in this setting, the first-best outcome can be achieved by taxing each product according to its marginal damages. For greenhouse gas emissions from automobiles, the tax on each car should be equal to the social cost of carbon times the car's average lifetime emissions, which is equal to average lifetime miles driven divided by fuel economy.[4] Absent the extensive margin, this is exactly what a gasoline tax would do, assuming that consumers are forward looking and rational.[5]

Alternative policies can be modeled as constraints on the planner's choice of the product tax vector. Our model uses a sufficient statistics approach to characterize the welfare loss of using any alternative set of taxes instead of the first-best tax schedule. We can then use that to model the effects of any particular policy that can be mapped into a tax schedule, which can be a linear function of one or more car attributes but also a step function or another non-linear scheme. For example, CAFE imposes an implicit tax schedule on vehicles that is a linear function of each vehicle's fuel consumption rate.[6] Our model can characterize the efficiency loss from using the second-best linear schedule instead of the first-best flexible one. This then indicates the difference in welfare between a gasoline tax and a CAFE-style policy.

### 2.1 Model setup

The model setup is as follows. A representative consumer purchases a portfolio of different types of goods, indexed $j = 1, ..., J$. For example, each $j$ would be a particular automobile model. The quantities chosen of each are denoted $x_j$. We assume that the set of models, and their characteristics,

[4] This setup allows for random failure rates. It is over the random failure of products that we take averages (expectations), rather than over different types of consumers.

[5] There is a literature that debates whether or not consumers fully value energy efficiency. Recent evidence is consistent with full valuation, or at most modest undervaluation (Busse, Knittel, and Zettelmeyer 2013; Allcott and Wozny 2014; Grigolon, Reynaert, and Verboven 2014; Sallee, West, and Fan 2015)

[6] Total fuel consumption for a given number of miles traveled is linearly related to per-mile fuel consumption, but non-linearly related to fuel economy. CAFE creates a shadow price linear in per-mile fuel consumption.

are fixed.

Consumers derive utility, $U$, from the consumption of these products: $U(x_1, ..., x_J)$. Each product generates an externality, denoted $\phi_j$, which we assume is linearly related to quantities and enters the social welfare function separably.[7] We denote the cost of production by $C(x_1, ..., x_J)$. There is an exogenous amount of income in the economy, $M$, and all remaining income is consumed in a quasi-linear numeraire, $n$. The planner maximizes social welfare, denoted $W$, which is just utility from the durable goods and numeraire, net of externalities. We write $W$ substituting the budget constraint in for the numeraire:

$$W = U(x_1, ..., x_J) + M - C(x_1, ..., x_J) - \sum_{j=1}^{J} \phi_j x_j. \tag{1}$$

We assume that there are many consumers, so that they do not internalize the externality when making their choice. And, we assume that the supply side is perfectly competitive. We assume that both $U(\cdot)$ and $C(\cdot)$ are twice differentiable and are weakly convex and concave, respectively.

The consumer perceives the prices of all goods to be constants, written as $p_j$, and pays taxes $t_j$. Thus, the consumer chooses a vector of goods to maximize:

$$\begin{aligned} \max_{x_1,...,x_J} \; & Z = U(x_1, ..., x_J) + n \\ \text{s.t.} \; & \sum_{j=1}^{J} (p_j + t_j) x_j + n \leq M. \end{aligned} \tag{2}$$

The consumer's first-order conditions imply that $\frac{\partial U}{\partial x_j} = (p_j + t_j)$. Under marginal cost pricing, this implies that $\frac{\partial U}{\partial x_j} - \frac{\partial C}{\partial x_j} = t_j$, which is just the standard tax wedge between marginal utility and marginal cost.

The planner can maximize utility by setting tax rates equal to external damages, that is, by choosing $t_j = \phi_j$ for all products. This simply imposes a Pigouvian tax on each good, and it fully internalizes the externality—the first-best is achieved. This can be seen by differentiating (1) with respect to each tax rate $t_j$:

$$\frac{dW}{dt_j} = \sum_{k=1}^{J} \left( \frac{\partial U}{\partial x_k} - \frac{\partial C}{\partial x_k} - \phi_k \right) \frac{\partial x_k}{\partial t_j} = \sum_{k=1}^{J} (t_k - \phi_k) \frac{\partial x_k}{\partial t_j}. \tag{3}$$

The first-order condition will be met ($\frac{dW}{dt_j} = 0$) when $t_j = \phi_j$.

Our aim is to characterize the welfare loss of deviating from this first-best schedule of Pigouvian taxes (in which $t_j = \phi_j \; \forall j$) to some arbitrary alternative tax schedule. We can then consider

[7] The $\phi$ terms can equally be thought of as the expected lifetime damages for a product that has a distribution of lifetime emissions, where uncertainty stems, for example, from random failure of a durable good. Heterogeneity in realized damages due to random failure will not influence the relative efficiency of product-based taxes as compared to fuel taxes, though it will result in different incidence on consumers who get above, or below, average lifetimes from their product.

particular alternative tax schedules, such as the second-best tax system where tax rates are a linear function of some other variable that is imperfectly related to damages, such as energy efficiency. To characterize the deadweight loss induced by an alternative tax schedule, we follow the sufficient statistics tradition of differentiating $W$ with respect to the tax and integrating. But, we have many tax rates and thus employ an intermediate algebraic step to derive our expression.

## 2.2 Derivation of sufficient statistics

Let any generic tax schedule be denoted as $\tau_1, ..., \tau_J$. We characterize the welfare loss of moving from the tax schedule $t_j = \phi_j$ to $t_j = \tau_j$ by specifying a weighted average of the two tax schedules and then integrating the marginal welfare losses of moving the weights from $\phi_j$ to $\tau_j$. That is, we specify the function $t_j = (1-\rho)\phi_j + \rho\tau_j$. We will differentiate $W$ with respect to $\rho$, and then characterize the welfare loss of moving from the optimal policy (when $\rho = 0$) to the alternative policy (when $\rho = 1$).

First, we differentiate equation (1) with respect to $\rho$ and substitute in the consumer's optimality condition. This yields:

$$\frac{dW}{d\rho} = \sum_{j=1}^{J}\sum_{k=1}^{J}\left(\frac{\partial U}{\partial x_j} - \frac{\partial C}{\partial x_j} - \phi_j\right)\frac{\partial x_j}{\partial t_k}\frac{\partial t_k}{\partial \rho} = \sum_{j=1}^{J}\sum_{k=1}^{J}(t_j - \phi_j)\frac{\partial x_j}{\partial t_k}\frac{\partial t_k}{\partial \rho}. \tag{4}$$

This term, $\frac{dW}{d\rho}$, is the incremental change in welfare as we move from the first-best rates toward the alternative tax schedule, where all rates move by an amount proportional to the difference between the first-best and the alternative taxes.[8] However, this object is not of particular interest to us; it is only an intermediate step that enables us to characterize deadweight loss in terms of demand derivatives (which are estimable) instead of the utility function (which is more difficult to recover with data).

Next note that, by our definition, $\frac{\partial t_k}{\partial \rho} = (\tau_k - \phi_k)$. This term, $(\tau_k - \phi_k)$, is the difference between the first-best and actual tax rates, which we call the "tax residual" for reasons that will become apparent. We use that substitution, as well as the definition of $t_j$, and simplify:

$$\begin{aligned}\frac{dW}{d\rho} &= \sum_{j=1}^{J}\sum_{k=1}^{J}(\{(1-\rho)\phi_j + \rho\tau_j\} - \phi_j)\frac{\partial x_j}{\partial t_k}(\tau_k - \phi_k)\\ &= \rho\sum_{j=1}^{J}\sum_{k=1}^{J}(\tau_j - \phi_j)\frac{\partial x_j}{\partial t_k}(\tau_k - \phi_k).\end{aligned} \tag{5}$$

Because $\rho$ is a constant, we can remove it from the summation, which yields the final equation. To get the change in social surplus from moving fully between the two tax schedules, we integrate

[8] This procedure is related to the technique developed in Hendren (2013), who labels an object that is similar to our $\frac{dW}{d\rho}$ the "policy elasticity".

from $\rho = 0$ to $\rho = 1$. If the demand derivatives are constant over the relevant range, then $\rho$ can be pulled out in front of the summation and the integration is straightforward. This yields:

$$W(\rho = 0) - W(\rho = 1) = -\frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}(\tau_j - \phi_j)(\tau_k - \phi_k)\frac{\partial x_j}{\partial t_k}. \tag{6}$$

This formula is in the form of a Harberger triangle; it is 1/2 times a tax wedge squared times a quantity derivative.[9] More specifically, it is closely related to a result in Harberger (1964), which characterizes the deadweight loss of taxing a good when distortions exist in multiple markets. The derivation in Harberger (1964) is different from ours, but the result is the same, except that our setup includes externalities whereas his does not (there are no $\phi$ terms in his result).

This expression can be decomposed to gain further insight. We denote $W(\rho = 1) - W(\rho = 0)$ as simply $DWL(\tau)$ and use the substitution $e_j \equiv (\tau_j - \phi_j)$ for the tax residuals. Rearranging equation (6) yields:

$$-2 \times DWL(\tau) = \underbrace{\sum_{j=1}^{J} e_j^2 \frac{\partial x_j}{\partial t_j}}_{\text{"own effects"}} + \underbrace{\sum_{j=1}^{J}\sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}_{\text{"cross effects"}}. \tag{7}$$

Equation (7) separates the deadweight loss formula into a term related to own-price derivatives (the way that a tax affects demand for the product itself) and a term related to cross-price derivatives (the way that a tax affects demand for alternative products). Under some conditions, which we discuss in detail below, the cross-price term will be zero, or at least small compared to the own-price term. We start with this special case to build intuition and then reintroduce the cross-price term.

When the cross-price term is zero, equation (7) simplifies to:

[9] This equation can alternatively be written in matrix notation. Note that

$$\frac{dW}{d\rho} = \begin{pmatrix} \tau_1 - \phi_1 & \dots & \tau_J - \phi_J \end{pmatrix} \begin{pmatrix} \frac{dx_1}{dt_1} & \dots & \frac{dx_1}{dt_J} \\ \dots & \dots & \dots \\ \frac{dx_J}{dt_1} & \dots & \frac{dx_J}{dt_J} \end{pmatrix} \begin{pmatrix} \tau_1 - \phi_1 \\ \dots \\ \tau_J - \phi_J \end{pmatrix} = (\tau - \phi)^T D(\tau - \phi),$$

where $D$ is the substitution matrix and $(\tau - \phi)$ is the vector of tax residuals. We can now rewrite

$$W(\rho = 0) - W(\rho = 1) = -\frac{1}{2}(\tau - \phi)^T D(\tau - \phi).$$

.

$$-2 \times DWL(\tau) = \sum_{j=1}^{J} e_j^2 \frac{\partial x_j}{\partial t_j} \tag{8}$$

$$= \underbrace{\sum_{j=1}^{J} e_j^2}_{\text{SSR}} \times \underbrace{\overline{\frac{\partial x_j}{\partial t_j}}}_{\text{mean derivative}} + \underbrace{\sum_{j=1}^{J} \left(e_j^2 - \overline{e^2}\right)\left(\frac{\partial x_j}{\partial t_j} - \overline{\frac{\partial x_j}{\partial t_j}}\right)}_{\text{covariance}}, \tag{9}$$

where bars indicate means of variables and $\overline{\frac{\partial x_j}{\partial t_j}} = J^{-1}\sum_{j=1}^{J} \frac{\partial x_j}{\partial t_j}$. This first line, equation (8), shows that the deadweight loss is determined by the magnitude of the errors in the tax rate and the partial derivatives of demand. The error in the tax rate $e_j$ is the tax wedge, relative to the socially efficient value. Thus, this formulation demonstrates that deadweight loss is the summation of a set of Harberger triangles across all the products. This is intuitive, as the assumption of the net zero cross-price terms implies that the sum of the partial equilibrium tax wedges are equal to the general equilibrium effect that considers all taxes.

Equation (9) provides additional insight by decomposing the Harberger triangle version into three parts, using the definition of the covariance. The first term is the sum of squared errors in the tax rate, multiplied by the average demand response (own-price derivative). We label the sum of squared errors as SSR because in an intuitive special case, discussed below, this will be precisely the sum of squared residuals from an OLS regression of lifetime damages $\phi_j$ on some observable metric upon which policy is based. The second term is the covariance between the squared errors in the tax rate and the own-price derivatives.

This formulation is particularly useful for understanding second-best linear tax schedules because it relates directly to ordinary least squares regression. Specifically, suppose that a planner is constrained to choose a tax schedule of the form $\tau_j = \alpha + \beta\theta_j$, where $\theta_j$ is some observable metric, such as a government energy efficiency rating. (CAFE, loosely defined, is such a policy.) The planner's second-best problem is then to choose $\alpha$ and $\beta$ to minimize deadweight loss (from equation (9)). When true damages $\phi_j$ are uncorrelated with own-price derivatives (so that the covariance term is zero), this second-best problem collapses to the standard OLS problem of minimizing residuals, where the residuals are the errors in the tax rates:

$$\min_{\alpha,\beta} = -\sum_{j=1}^{J} \frac{\partial x_j}{\partial t_j} \times \sum_{j=1}^{J} (\phi_j - \alpha - \beta\theta_j)^2. \tag{10}$$

This is the standard OLS optimization problem, and the solution will be the standard OLS regression coefficients. In that case, the sum of squared errors in the tax schedule will be exactly the sum of squared residuals from an OLS regression of lifetime damages $\phi_j$ on $\theta_j$, which is the metric that determines policy treatment.

For example, $\theta_j$ might be official fuel consumption (inverse of fuel economy) ratings. As discussed above, CAFE implicitly imposes a set of taxes and subsidies for vehicles that are a linear function of fuel consumption.[10] If there were no heterogeneity in average lifetime mileage of different models of cars, then this linear function could exactly predict lifetime damages; i.e., the $R^2$ of a regression of lifetime fuel consumption on fuel consumption ratings would be 1, and the sum of squared residuals would be zero. To the extent that there is utilization heterogeneity across models with the same fuel efficiency, however, the observable metric (fuel consumption) will be imperfectly correlated with true damages. The deadweight loss of the second-best linear policy is then directly tied to the sum of squared residuals from the linear regression, as captured in equation (10). Intuitively, the deadweight loss from these imperfect tax rates are scaled by the size of own-price demand derivatives. As demand is more sensitive to prices, the inefficiency of getting prices wrong rises.

These equations express deadweight loss in dollars, but we can also express the welfare gain of using the optimal second-best linear policy (the OLS fitted line for the special case where damages are uncorrelated with own-price derivatives), in lieu of an efficient policy. To do so, we need some benchmark policy. One useful benchmark is a policy that puts a single tax rate on all products. This policy does not change the relative price of any of the products, but it allows the policymaker to impose the correct average tax on the extensive margin to correct the size of the market. In terms of equation (10), this baseline policy allows the planner to choose only a constant term, but forces the slope coefficient to be zero.

Compared to this baseline, the efficiency gain of using the linear policy over the baseline, divided by the efficiency gain of the first-best policy over the same baseline is equal to the $R^2$ from the regression:

$$\frac{\text{DWL(Linear)} - \text{DWL(Constant)}}{\text{DWL(First best)} - \text{DWL(Constant)}} = \frac{ESS}{TSS} = R^2. \tag{11}$$

This formulation is particularly useful because it relates directly to data. In Section 4 we demonstrate this for the case of vehicle fuel economy policy. There, we estimate the total lifetime damages of different vehicle models using data on lifetime mileage shortly before vehicles are scrapped. We then calculate the degree to which a CAFE-style policy is inefficient by estimating the best linear fit and calculating the $R^2$. We then use estimates of the social cost of carbon and of the own-price demand elasticity of cars, which we take from the literature, to convert the $R^2$, which characterizes welfare in proportional terms, into dollar amounts.

Actual policies, of course, need not be second-best. The formula in equation (7) is a valid equation for any alternative policy, provided that demand derivatives are constant over the relevant range of tax rates. (When demand derivatives are not constant, the formula is an approximation, which will be a good approximation for "small" changes in the way that traditional Harberger

[10] Note that CAFE is additionally constrained on the extensive margin because vehicles with fuel-economy ratings larger than the standard be are subsidized, while vehicles with fuel-economy ratings below the standard are taxed. Thus, CAFE will fail to shrink the overall size of the car market by the optimal amount, as would a gas tax. Our characterization of CAFE as a linear tax schedule abstracts from this "overall market size" issue.

triangles are.) One useful decomposition is to separate the mean bias in tax rates from their variance, which can be seen by rewriting deadweight loss for the special case with zero net cross effects and no correlation between tax rate errors and demand derivatives, as follows:

$$-2 \times DWL(\tau) = \sum_{j=1}^{J} \frac{\partial x_j}{\partial t_j} \times \left( \underbrace{J \times \bar{e}^2}_{\text{bias}} + \underbrace{\sum_{j=1}^{J} (e_j - \bar{e})^2}_{\text{variance}} \right). \tag{12}$$

This illustrates that there is a bias in the tax rates, and there is a variance in the tax rate errors, and their effects on welfare can be separated. Besides a mean bias, a linear policy can also have a "slope bias": the slope of the policy differs from the slope estimated by OLS. For instance, linear taxes will have a slope bias if there is a correlation between lifetime utilization and energy efficiency ratings that is not accounted for in policy design. The mean bias can be eliminated by a linear policy, but the variance cannot. OLS minimizes the variance; policies with a slope bias have a larger variance term.

### 2.3 When are cross-price effects small?

The above discussion focused on the case when the cross-price effects in equation (7) were zero. The intuition for this can be seen by focusing on cases where the tax rates are unbiased, that is, where $\bar{e_j} = 0$. In that case, the cross-price term is the sum of the product of three things, cross-price demand derivatives (which will generally be positive) and two different tax rate errors, $e_j$ and $e_k$. In the unbiased case, the errors are zero on average, so, on average, the cross-price terms are multiplying $e_j$ and $e_k$, and these products will be sometimes positive, sometimes negative. As a result, it is possible for the cross-price effects to cancel out. In contrast, so long as all own-price derivatives are negative (i.e., there are no Giffen goods), the own-price terms in equation (7) must each be positive, so the summation will be positive.

Again, the OLS case helps with intuition. The sum of the product of all pairwise combinations of residuals from OLS is zero. Thus, when the alternative tax schedule $\tau$ is the OLS best fit, the cross-price term could be zero. If the tax rates are biased, then the cross-price effects will not be zero on average, but will have a bias term.

The cross-price effects also involve the cross-price derivatives, so the OLS residual intuition is not sufficient to tell us when this summation will net to zero. To see this further, we can decompose the cross-price effects further to characterize them as a function of covariance and bias.

$$-2 \times DWL(\tau) = \underbrace{\sum_{j=1}^{J} e_j^2 \frac{\partial x_j}{\partial t_j}}_{\text{“own effects”}} + \sum_{j=1}^{J} e_j \times \left\{ \underbrace{\sum_{k \neq j} \left( (e_k - \overline{e_k}) \left( \frac{\partial x_j}{\partial t_k} - \overline{\frac{\partial x_j}{\partial t_k}} \right) \right)}_{\text{“covariance term”}} + \underbrace{(J-1) \times \overline{e_k} \overline{\frac{\partial x_j}{\partial t_k}}}_{\text{“bias term”}} \right\}, \quad (13)$$

where $\overline{\frac{\partial x_j}{\partial t_k}} = (J-1)^{-1} \sum_{k \neq j} \frac{\partial x_j}{\partial t_k}$ and $\bar{e_k} = \sum_{k \neq j} e_k$, which is defined separately for each $j$. For the unbiased policy, the bias term in the bracketed expression will be zero, leaving only the covariance term. The covariance term is the covariance between the error in the tax rate and the degree of substitutability across alternative products $k$ for a given model $j$. This will be zero, on average, when the errors in the tax rates are uncorrelated with the degree of substitutability across products.

We do not expect that the cross-price terms will be exactly zero in real settings for second-best policies because products with similar differences in utilization from the average are also likely to be closer substitutes than average. For example, if luxury vehicles tend to last longer than average, and also tend to be close substitutes for each other, we would expect the cross-price effects overall to be nonzero.

Even when the cross-price terms are not zero it is still intuitive to expect that they may be small, as compared to the own-price terms, because they have the tendency of canceling each other out (the sum of the product of all pairwise combinations of tax residuals $e_j e_k$ is zero). But, this is ultimately an empirical question. We thus emphasize the theoretical intuition of the special case with zero cross-price terms here and then, in our empirical investigation below, determine the degree to which cross-price effects are likely to be important by modeling a range of possible correlations between tax rate errors and demand derivatives.

# 3 Data

Our empirical analysis considers the case of automobiles. It uses data on vehicle miles traveled from the emissions compliance program in California matched to a comprehensive registration micro data that allows us to infer when a vehicle has been retired. Our analysis is primarily based upon the universe of emissions inspections from 1996 to 2010 from California's vehicle emissions testing program—the Smog Check Program, which is administered by the California Bureau of Automotive Repair (BAR). An automobile appears in the data for a number of reasons. First, vehicles more than four years old must pass a smog check within 90 days of any change in ownership. Second, in large parts of the state an emissions inspection is required every other year as a pre-requisite for renewing the registration on a vehicle that is six years or older. Third, a test is required if a vehicle moves to California from out-of-state. Vehicles that fail an inspection must be repaired and receive another inspection before they can be registered and driven in the state. There is also a group of exempt vehicles. These are: vehicles of 1975 model-year or older, hybrid and electric vehicles, motorcycles, diesel-powered vehicles, and large natural-gas powered trucks.

These data report the location of the test, the unique vehicle identification number (VIN), odometer reading, the reason for the test, and test results. We decode the VIN to obtain each vehicle's make, model, engine, and transmission. Using this information, we match the vehicles to EPA data on fuel economy. Because the VIN decoding is only feasible for vehicles made after 1981, our data are restricted to these models. This yields roughly 120 million observations.

The primary use of the smog check data is to calculate the vehicle's odometer reading at the time the vehicle was scrapped. However, vehicles may leave the smog check data because they leave California. To accurately measure when a vehicle is scrapped, we also use data obtained from CARFAX Inc. These data contain the date and location of the last record of the vehicle, regardless of state, reported to CARFAX for 32 million vehicles in the smog check data.[11] Because the CARFAX data include import/export records, we are able to correctly classify the outcomes of vehicles which are exported to Mexico as censored, rather than scrapped, thus avoiding the issues identified in Davis and Kahn (2010). We define a vehicle as being scrapped if the vehicle is not registered anywhere in the US for two years. The CARFAX data do not report odometer readings; therefore, we restrict analysis to vehicles that were in the smog check program two years prior to being scrapped.

Together, the smog check and CARFAX data give us a measure of the total lifetime vehicle miles traveled (VMT) for a particular unit, but we do not observe all units, which creates the possibility of statistical bias. The $R^2$ in an OLS regression, which is our primary statistic of interest, is biased downwards if there is classical mismeasurement in the dependent variable (lifetime mileage), which could arise in our case from sampling variation.[12] We have a large sample, but we only observe a limited number of retirements for many vintages of less popular models. This can cause a downward bias in the $R^2$ of a regression of lifetime gasoline consumption on fuel economy, but in Section 4 we report evidence that the bias appears to be very small in economic terms.

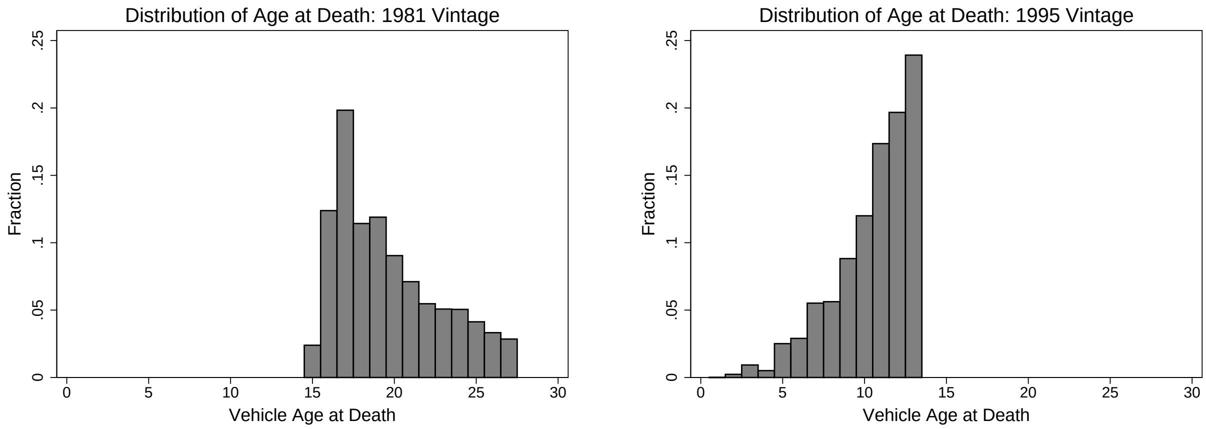
In addition, we observe relatively few vehicles that are under six years old because most of them are not required to take an emissions test. And, we do not observe the final odometer reading for vehicles that have not yet been retired. This creates a censoring problem, which will vary across model years. To illustrate this, Figure 2 shows the age at retirement of vehicles that appear in our sample for model year 1981 and 1995 vehicles separately. Because our data on retired vehicles span the period from 1996 to 2008, we observe 1981 vehicles that were at least 15 years old at retirement, whereas we observe retirements up to age 13 for 1995 models.[13] This censoring can create (non

[11] The actual date of retirement of the vehicle is not the same as the last date of registration. The vehicle's odometer reading occurs at the last registration date. Rather than imputing the odometer at the moment of scrap using hazard rates, we simply use the last observed reading for reasons of transparency. Such an imputation would be unlikely to have an impact on the $R^2$ in our regressions.

[12] When the independent variable is correlated with the true dependent variable but not the white noise mismeasurement of it, the $R^2$ must be driven down, and more so as the variance in the noise is large relative to the variance in the true dependent variable. Intuitively, this is because noise inflates the total variation in the observed outcome, but it leaves the explained variation unchanged.

[13] Our smog check data extend to 2010, but we must observe a two year window after a vehicle's last smog check

**Figure 2:** The Distribution of Vehicle Age at Death for Different Vintages




classical) mismeasurement, which will be particularly problematic when comparing across cohorts.

In this draft, we explore the importance of censoring for our empirical conclusions by comparing how our results vary when we use all of our data versus when we restrict attention to model years 1988 to 1992, for which censoring will be less problematic. We also use additional data and extrapolation techniques to paint a more complete picture of lifetime VMT for each vehicle in our sample. Specifically, we have access to comprehensive national registration data from R.L. Polk. These data measure the stock of registered vehicles of each 10-digit VIN-prefix ("VIN10-prefix") in each year. Thus, we can use these data to measure the number of vehicles that are censored and to determine the scrappage rate of models for certain years (and ages) outside of our smog sheck sample. We will use this additional information, along with the pattern of mileage of related models in the data to predict lifetime VMT for all censored vehicles. Our empirical analysis uses these extrapolated data to provide an upper bound on $R^2$. We are working on alternative extrapolation techniques to further explore the robustness of our $R^2$ estimates to censoring. In Section 4, we show robustness of our results across various samples and conclude that the censoring may ultimately have a limited impact on our conclusions.

# 4 Numerical Results for the Representative Consumer Model

In this section, we use our theoretical results and our automobile data to quantify the deadweight loss that results from using second-best policies. We focus first on the $R^2$ of regressions relating fuel-economy rating to lifetime gasoline consumption. We report the $R^2$ measure from a variety of specifications which take alternative approaches to dealing with the limitations of our data, namely censoring problems and sampling error. As discussed in Section 2, under certain simplifying assumptions, the $R^2$ alone is a sufficient statistic. Specifically, it is equal to the fraction of the

to know if it has missed its next required check. Thus, we identify vehicle retirements that occurred between 1996 and 2008.

welfare gain achieved by the first-best (over a baseline policy that places a common tax rate on all automobiles) that can be achieved by the optimal second-best linear policy; this fraction is an upper bound on the welfare gain from CAFE.[14] Other factors will enter the calculation that determines economic efficiency when our simplifying assumptions do not hold, or when we wish to evaluate a policy other than the second-best linear one. In those cases, the $R^2$ will remain a rough indicator of economic efficiency. In ongoing work, we are using simulations to determine how important these other factors may be. In Section 4.2, we use estimates of the social cost of carbon and the derivative of vehicle demand with respect to price to convert the $R^2$ into deadweight loss measured in dollars.

Our empirical analysis makes several important assumptions. First, we focus here on greenhouse gas emissions, in part because our data allow us to look at this externality more easily than others. Our data indicate the total miles that a vehicle has been driven, but they do not tell us *where* those miles were driven. This is not a problem for global pollutants like carbon dioxide, but it means that we can paint only an incomplete picture of the heterogeneity in damages from local air pollution. Our focus on greenhouse gases likely understates the overall importance of heterogeneity by omitting local air pollution, but it is conceivable that the two sources of heterogeneity could partially offset. We intend to incorporate local pollutants in the future.[15]

Second, we assume that in-use fuel economy matches the EPA's estimate of the combined fuel economy rating, for each type of vehicle on average across drivers and miles driven. If some vehicle models are driven disproportionately in city traffic, while others are driven disproportionately on highways, or if the owners of some vehicles drive particularly aggressively, then the differences between laboratory test ratings and average in-use performance will differ across models. Either factor could be substantial, and by omitting such factors we suspect that we are understating heterogeneity. But, it is possible that heterogeneity in the deviation between on-road fuel economy and official test ratings is negatively correlated with durability, thereby offsetting some of our measured heterogeneity. We will explore this issue in the future.[16]

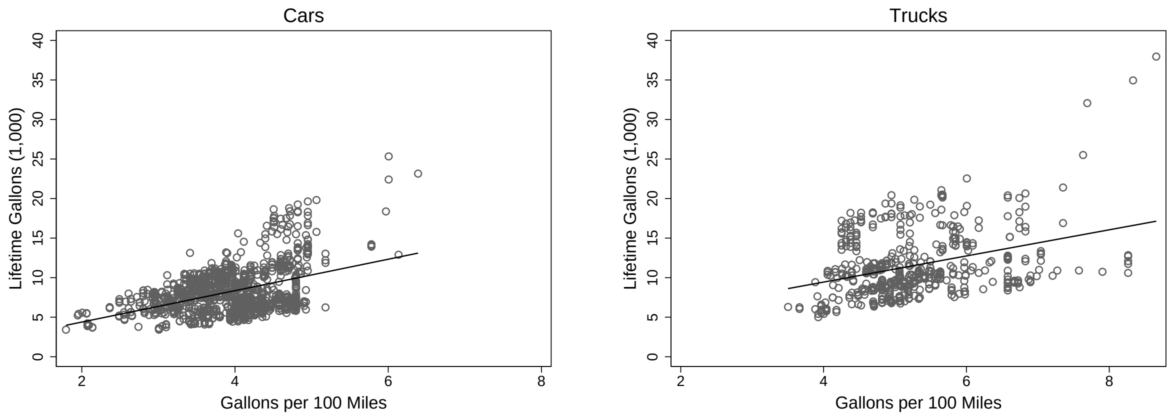
Third, at present, we ignore the *timing* of emissions. That is, we sum total miles driven and do not discount them into the present value at the time when a car is new. We do so both for simplicity, but we also note that the time path of the social cost of carbon suggested by current

[14] Because a fuel-economy standard will have the wrong intercept, Holland, Hughes, and Knittel (2009) show that there is no guarantee that welfare will increase, relative to the case of no regulation.

[15] Registration data and survey data could allow us to identify the population density and local air quality of the location where each type of vehicle is registered. We are currently exploring the possibility of estimating heterogeneity in local air pollution by vehicle type using these data sources. See Knittel and Sandler (2013) for a related analysis of how heterogeneity in local air pollution across vehicles influences the efficiency of a fuel tax as a policy instrument.

[16] We can use a combination of registration data and survey data from the National Household Transportation Survey to assess the degree to which vehicles are registered and driven in rural versus highway settings. This analysis will have limitations, but it will likely allow us to make an order of magnitude determination about this source of heterogeneity. Regarding heterogeneity due to driving style, see Langer and McRae (2014) for evidence that the variation across consumers in on-road fuel economy is vast. Note that what matters for our analysis is not the variance across individuals, but the degree to which the average consumer of a particular model differs from the average consumer of another model. Their analysis is based on data from a single type of vehicle, so it offers little guidance for our purposes.

**Figure 3:** The Relationship Between Lifetime Gasoline Consumption and Fuel-Efficiency




Note: The unit of observation is a type of vehicle (a VIN10-prefix), and gallons consumed is the average across observations for that type. The sample is restricted to models for which we observe at least 200 vehicle retirements from model years 1988 to 1992. Observations with VMT $\geq 1,000,000$ miles are dropped. Solid lines are OLS prediction lines.

federal guidelines roughly offsets discounting. That is, as the social cost of carbon is set to rise at roughly the rate of interest, accounting for the time path of the social cost of carbon and the time path of vehicle emissions will yield a result that is quite similar to a baseline that ignores discounting and uses the current social cost of carbon and full lifetime mileage without discounting.

Fourth, at present, we ignore carbon emissions stemming from the construction (and scrappage) of each vehicle. We can account for this in the future by estimating the carbon emissions "embodied" in the manufacturing process.

Fifth and finally, before proceeding to the results, we remind the reader that our welfare analysis is derived in a simplified model that does not allow for lifetime VMT to respond to policy (no intensive margin response) and assumes that product attributes are fixed (i.e., we consider the short run).

## 4.1 The Degree of Heterogeneity in Durability Across Models

To begin, we show scatterplots of the relationship between lifetime gasoline consumption and official fuel consumption ratings. Figure 3 shows the relationship between a model's total lifetime externality (gallons of gasoline) and its fuel economy, for both cars and trucks. A point in the figure corresponds to the average gasoline consumption at the VIN10-prefix level, which corresponds roughly to a unique make, model, vintage, and engine size, and sometimes also distinguishes trim levels and transmissions. The figure ignores within-VIN10 variation in gasoline consumption. The sample in the figure is restricted to model years 1988 to 1992, the years for which censoring is least problematic (we observe retirements between 1996 and 2008), and to vehicle models for which we have at least 200 observed retirements. This mitigates sampling error and reduces the number of observations for visual clarity. In addition, we drop observations above 1,000,000 miles to limit the influence of outliers. The figures also show an OLS fitted line for reference.

The figure illustrates that there is, as expected, a strong, positive correlation between fuel consumption ratings (the inverse of fuel economy ratings) and lifetime gasoline consumption. But, there is also a great deal of dispersion. Vehicles have significantly different average longevity (total lifetime mileage), and this translates into variation in lifetime fuel consumption, conditional on the official fuel consumption rating. The $R^2$ for cars and trucks in this sample is only 0.18 and 0.12, respectively. (The $R^2$ from a combined sample regression is 0.29.) When the assumptions hold for the special case of our model, this implies that the second-best linear policy captures only 18% and 12% of the welfare gains for cars and trucks that would be achievable with an efficient set of product-based taxes that varies not only with fuel economy, but also with vehicle durability.

The figures are drawn with a particular subset of the data. To illustrate how the implied efficiency of second-best policies varies with different sample restrictions, Table 1 reports the $R^2$ from a set of regressions that take the form:

$$\text{Average Lifetime Gasoline Consumption}_j = \alpha + \beta \text{Gallons per Mile}_j + \varepsilon_j, \tag{14}$$

where $j$ indexes a vehicle type (VIN10-prefix). We report a range of estimates in order to assess the importance of sample restrictions, weighting, censoring and sampling error.

Table 1 shows our estimates of the $R^2$ for different sample restrictions, both using OLS and weighted least squares (WLS), where VIN-prefixes are weighted by the number of observed retirements $N$. In all cases, we drop observations with reported mileage above one million (1,525 observations out of roughly 4 million, or less than 0.05%). The first two panels treat a VIN10-prefix or VIN8-prefix (which groups model years together, but still distinguishes make, model, engine type, etc.) as a unit of observation, consistent with Figure 3 above. The third and fourth panels use the microdata: the unit of observation is a retired vehicle.

Panel 1 shows that our estimate of the $R^2$ remains small in all VIN10-prefix specifications, ranging from a low of 0.17 to a high of 0.29. $R^2$ is slightly higher when the data is collapsed at the VIN8-prefix level (0.19 to 0.34).

Importantly, our estimates change very little when we restrict the sample to include only 1988 to 1992 model years (panel 2). As these model years span the age range in which the majority of retirement happens, this provides us with a first indication that our welfare conclusions will be broadly robust to additional measures that account for censoring in the data.

As discussed above, white noise in the measurement of lifetime mileage by type (sampling error), should cause a downward bias in the $R^2$. To assess the importance of sampling error, we compare results from OLS to WLS, which weights models by the number of vehicles that is observed as being scrapped. We also check how our results change when we limit the sample to vehicles for which we observe relatively many retirements ($N \geq 200$). The $R^2$ changes only modestly when moving between OLS and WLS, and when restricting the sample to $N \geq 200$. This suggests that our qualitative findings are not overly sensitive to sampling considerations. We explore this issue

**Table 1:** Regression $R^2$ Using Raw Data

| **Regressions using VIN-prefix averages** | | | | |
|---|---|---|---|---|
| | **VIN10-prefix** | | **VIN8-prefix** | |
| | **OLS** | **WLS** | **OLS** | **WLS** |
| **All model years** | | | | |
| All models | .26 | .20 | .23 | .19 |
| Models with $N \geq 200$ | .22 | .17 | .27 | .19 |
| **Model years 1988-1992** | | | | |
| All models | .27 | .26 | .28 | .27 |
| Models with $N \geq 200$ | .29 | .22 | .34 | .25 |
| **Regressions using microdata** | | | | |
| **All model years** | | | | |
| All models | .08 | | | |
| **Model years 1988-1992** | | | | |
| All models | .10 | | | |

Note: Table shows $R^2$ from regressions of average lifetime gallons consumed on fuel consumption rating. The unit of observation is either a VIN10-prefix or a VIN8-prefix, except for the last two rows, which include individual vehicles. Observations with $VMT \geq 1,000,000$ miles are dropped. $N$ is the number of observed retirements $N$, and WLS weights the regressions by $N$.

further below.

For reference, panels 3-4 also report the $R^2$ from the OLS regression on our underlying microdata, rather than on the data collapsed to VIN-prefix averages. The $R^2$ is 0.08 for all model years and 0.10 for model years 1988-1992. It is important to emphasize that this is *not* the relevant measure in the representative consumer model, as these regressions include heterogeneity in mileage across different individual drivers of the same vehicle model. As such, it includes differences in how individual drivers depreciate their vehicles, including accidents that lead vehicles to be scrapped. Accidents are *ex post* realizations of random product failure. As discussed in the introduction, such random failures will not influence the relative efficiency of one policy versus another and is therefore not the prime object of our study. We include it here to demonstrate the full degree of heterogeneity in the underlying microdata.

Overall, the relatively low $R^2$ statistics suggest that there is substantial variation in lifetime gasoline consumption, conditional on fuel-economy ratings. We have also explored a number of alternative estimates that treat cars and trucks separately and that estimate the $R^2$ for each model year separately. In all cases, the qualitative conclusion remains that there is substantial variation in lifetime consumption that is not explained by fuel economy, which implies that policies based only on fuel economy ratings, not on average product durability, will raise welfare by significantly less than would an efficient policy (including a carbon tax or a gasoline tax).

#### 4.1.1 Policies May Also Be "Biased"

Our approach can be altered to consider something other than just the second-best linear policy. For a nonlinear fuel economy policy, the $R^2$ would still be the relevant summary statistic, but the relevant independent variable would be different. For example, if the fuel economy policy were more flexible so that it could put a shadow price on each model that was a quadratic function of fuel consumption ratings, we could estimate the $R^2$ from a quadratic regression. If the policy determines tax rates based on fuel economy bins or a threshold fuel economy rating, the $R^2$ from a regression with dummy variables would be the relevant statistic.

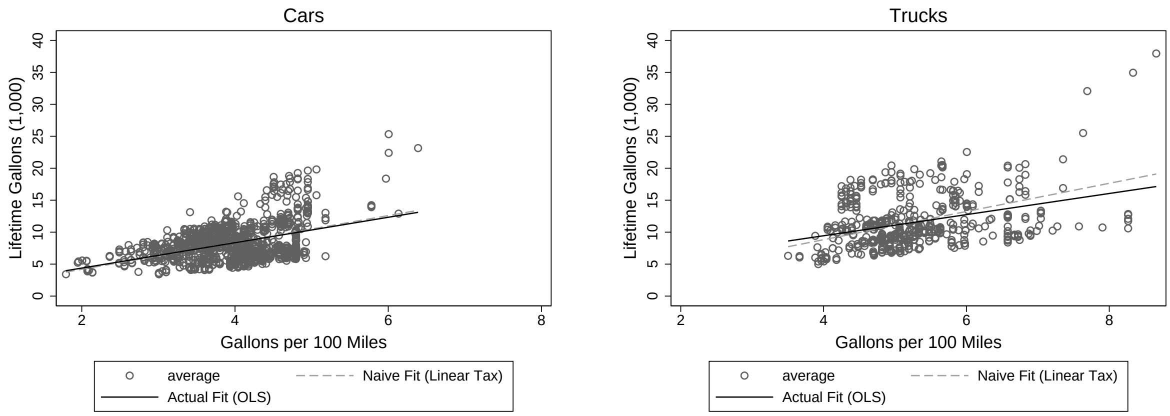
Alternatively, if the linear policy used is not the second-best (that is, if the predicted line is biased), then this bias causes an additional source of welfare loss, as demonstrated in equation (12). A bias in the mean tax rate fails to shrink the car market by the optimal amount (this is only relevant in a model with an outside good). A bias in the slope of the linear tax schedule will increase tax errors. This will occur when there is a correlation between average lifetime mileage and fuel consumption ratings in the data, but the policy is determined as if there is no such correlation. We illustrate this in Figure 4, which replicates Figure 3, but adds a line that represents the relationship between fuel consumption ratings and lifetime fuel consumption, if all cars (or trucks) were driven the same number of miles, which we set equal to the observed mean in our data. This line represents the best fit line that a policymaker would choose if they knew only the average mileage (separately for cars and trucks) across all vehicles, but did not know the correlation between average mileage and fuel consumption ratings. This is our depiction of a "naïve" linear tax, which gets the average shadow price right, but ignores durability completely. The current CAFE standards are naïve in this way, as the standards are not based on expected VMT. Figure 4 shows that the naïve linear tax differs noticeably from the best linear tax for trucks, but that the difference for cars is small. This mispricing represents another source of inefficiency from ignoring heterogeneity in durability. It turns out not to be the dominant concern in our data, but it may be important in other contexts.

#### 4.1.2 Additional Robustness Checks: Outliers, Sampling and Censoring

Our data include some cases of very high lifetime VMT, which raises the possibility of coding errors. Our estimates of the $R^2$ could be sensitive to such outliers, even when restricting to vehicles with relatively large sample sizes. In the results above, we have dropped observations for which VMT-at-death exceeds 1,000,000 miles. To check whether our $R^2$ results are sensitive to this sample restriction, we have run regressions that include all observations as well as regressions in which we winsorize the underlying micro data at different VMT thresholds. When we include the vehicles with the highest mileage, $R^2$ is virtually unchanged. The $R^2$ increases only modestly when we winsorize the highest VMT vehicles at increasingly stringent levels. The OLS $R^2$ rises from a baseline of 0.28 to a maximum of 0.37 when we limit the influence of data over 400,000 miles. The WLS $R^2$ rises from a baseline of 0.22 to a maximum of 0.30 for the same restriction. See Appendix Table A.1 for details. Our qualitative conclusions are therefore robust to outliers.

Winsorizing the data at a level as low as 400,000 miles does seem to be restrictive; that is,

**Figure 4:** The Relationship Between Lifetime Gasoline Consumption and Fuel-Efficiency



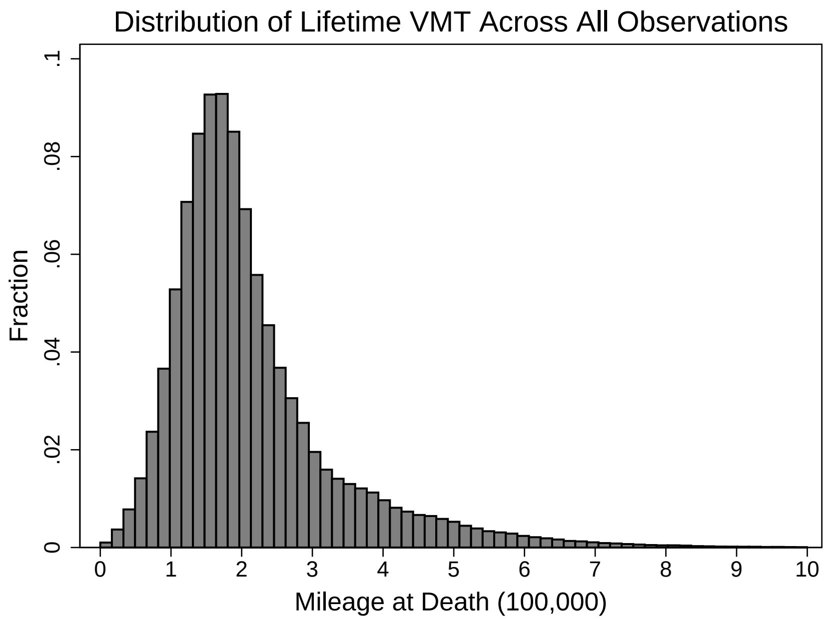

Note: The unit of observation is a type of vehicle (VIN10-prefix), and gallons consumed is the average across observations for that type. The sample is restricted to models for which we observe at least 200 vehicle retirements from model years 1988 to 1992. Observations with $VMT \geq 1,000,000$ miles are dropped. Solid lines are OLS prediction lines. Dashed lines are linear fits under the assumption that all vehicles are driven the mean number of miles.

we suspect that most of the cases of reported high lifetime VMT are legitimate. To demonstrate this, we plot the full histogram of lifetime odometer readings across all of our microdata in Figure 5. The data have a mode around 160,000 miles, but there is a long right tail. Just under 7% of vehicles in our data are scrapped with over 400,000 miles. It is useful to recall that our data are for California, where the climate may facilitate longer vehicle lifetimes than would be true in other climate zones.

Above we argued that bias in the $R^2$ due to mismeasurement from sampling variation was likely to be small because our results are not overly sensitive to restricting the set of vehicles to those with a large sample. To further examine the importance of sampling variation, we test how the $R^2$ changes when we randomly select subsets of our data for analysis. Specifically, we limit our sample to all VIN10-prefixes for our focal vintages of 1988 to 1992, for which we have at least 200 retirements in our sample. We then bootstrap that sample and estimate the $R^2$ many times. The mean estimate is 0.283 (which corresponds to the parallel specification in 1). Next, we bootstrap the sample again, but in each iteration we drop 50%, 90% or 98% of our sample randomly. Dropping these fractions of the sample decreases the $R^2$ to 0.282, 0.273 and 0.229, respectively. The negligible change in $R^2$ as the sample size is cut in half provides strong evidence that sampling error is unlikely to cause a downward bias in our $R^2$ estimates. Even cutting our data down to just 2% of our preferred sample reduces the $R^2$ by only 20 percent.

Finally, we further assess the importance of censoring. Table 1 above showed that restricting the sample to model years 1988-1992 does not affect the $R^2$ much, providing a first indication that the bias from censoring may be limited. Here we consider two alternative methods. The first method is an extrapolation technique that assigns retirement counts and VMT-at-death to non-observed

**Figure 5:** Distribution of Lifetime VMT Across All Observations




ages for each individual VIN10-prefix. The extrapolation is intentionally conservative, so that the resulting $R^2$ should be considered an upper bound on the true $R^2$. The second method exacerbates the censoring by progressively removing vehicles of certain ages, and shows how the $R^2$ changes in response.

The extrapolation method starts with the national registration data at the VIN10-prefix level. We first compute annual scrap rates for each VIN10-prefix over the sample period 1999-2009 and fill in missing scrap rates wherever possible. Next, we fill in missing scrap rates for unobserved ages using average scrap rates by age at the VIN8-prefix level, which does not distinguish model year. In other words, if the scrap rate for a 20-year-old 1985 Toyota Corolla LE is missing, we replace it with the average scrap rate of any 20-year-old Toyota Corolla LE, regardless of vintage (assuming that at least one vintage is observed at age 20). For ages that are not observed at the VIN8-prefix level, we assign scrap rates based on sample-wide average scrap rates by age (weighted by registration counts).[17] Having extrapolated missing scrap rates (and, indirectly, missing vehicle retirements), we then impute missing VMT-at-death using a similar procedure. We first replace VMT-at-death for each age without data using VMT averages across VIN8-prefixes. For ages that are never observed at the VIN8-prefix level, we use a similar polynomial fit for the relationship between VMT-at-death and age, averaged across all models and weighted by the number of retirements.

This is an extremely conservative approach, in that we assume that missing scrap rates and VMT-at-death are the same across *all* vehicles. This is "conservative" in that it necessarily reduces cross-model variation in lifetime mileage and thus raises the $R^2$. The process essentially removes

[17]Specifically, we fit a fifth-order polynomial to the scrap rate by age pattern, and use the polynomial fit for imputing missing data.

all relevant variation for the imputed observations. The resulting $R^2$ from regressions with imputed data should therefore be considered an upper bound, one that is likely substantially above the true $R^2$ that would be obtained with a fully uncensored sample.

**Table 2:** Regression $R^2$ Using Imputed Data

| **VIN10-prefix averages, model years 1988-1992** | **OLS** | **WLS** |
|---|---|---|
| **VMT imputed for all models** | | |
| All models | .44 | .43 |
| Models with $N \geq 200$ | .45 | .38 |
| Models with $N_{imputed} \geq 400$ | .47 | .40 |
| **Only models for which VMT is imputed for $\leq 12$ ages** | | |
| All models | .34 | .25 |
| Models with $N \geq 200$ | .29 | .24 |
| Models with $N_{imputed} \geq 400$ | .28 | .23 |

Note: Table shows $R^2$ from regressions of average lifetime gallons consumed on fuel consumption rating, where scrap rates and VMT for missing ages are imputed. The unit of observation is a VIN10-prefix. Observations with $VMT \geq 1,000,000$ miles are dropped. WLS uses the actual number of observed retirements $N$ when the sample is selected based on $N \geq 200$ and the imputed number of retirements $N_{imputed}$ when the sample is selected based on $N_{imputed} \geq 400$.

Table 2 presents the results for model years 1988-1992 using this imputation. When missing data are imputed for all models, the $R^2$ increases to 0.38-0.47, depending on whether the regression is weighted and if the sample is restricted to observations with at least 200 observed retirements or at least 400 imputed retirements (panel 1). While this range is clearly above 0.22-0.29 (as reported in panel 2 of Table 1), the $R^2$s are still low from an absolute perspective. Panel 2 shows that when we restrict the sample to VIN10-prefixes for which we impute VMT-at-death for at most 12 ages, the range goes down to 0.23-0.34. This provides further evidence that censoring is unlikely to cause a large bias in our results.

Our second approach to investigating the impact of censoring is to drop vehicles of certain ages, thereby exacerbating the censoring problem, to see how that influences the $R^2$. The idea is that the change in $R^2$ in response to more restrictive censoring can provide additional insight into what would happen if we could instead relax the censoring.

Specifically, we restrict the sample to models with $N \geq 200$ and model years 1988-1992 and show how the $R^2$ estimates change as we progressively remove vehicles of older ages from the sample. Table 3 shows the results for vehicles in the age ranges 10-$X$ years old, where $X$ goes up from 11 to 20 years. We find that the $R^2$ increases from 0.28 to 0.40 when the age range is further censored, suggesting that *less* censoring would yield lower values. We have also run age-specific regressions (i.e., regressions on only 10-year-old cars). The $R^2$ falls as vehicles get older. Intuitively, it seems that censoring "young" vehicles depresses the $R^2$, as there will be less variation in VMT among cars that are scrapped (generally because of accidents) at young ages, whereas censoring "old" vehicles likely exaggerates the $R^2$ by understating heterogeneity.[18] The smog check data are censored to omit vehicle deaths below six years, but relatively few vehicle deaths occur in those years, so on balance our data are mostly missing deaths at older ages. This suggests that the censoring problem is most likely, on net, causing us to exaggerate the $R^2$.

**Table 3:** Regression $R^2$ With Different Vehicle Age Restrictions

| **VIN10-prefix averages, model years 1988-1992,** $N \geq 200$ | | | |
|---|---|---|---|
| **Low age** | **High age** | **OLS** | **WLS** |
| 10 | 10 | .37 | .41 |
| 10 | 11 | .40 | .37 |
| 10 | 12 | .39 | .34 |
| 10 | 13 | .37 | .32 |
| 10 | 14 | .35 | .31 |
| 10 | 15 | .33 | .29 |
| 10 | 16 | .31 | .27 |
| 10 | 17 | .30 | .25 |
| 10 | 18 | .29 | .23 |
| 10 | 19 | .29 | .22 |
| 10 | 20 | .28 | .21 |

Note: Table shows $R^2$ from regressions of average lifetime gallons consumed on fuel consumption rating. The unit of observation is a VIN10-prefix. Observations with $VMT \geq 1,000,000$ miles are dropped. WLS weights the regressions by the number of observed retirements $N$.

## 4.2 Estimates of Deadweight Loss

We can translate the relative gains from the first- and second-best product-based taxes, expressed above as an $R^2$, into deadweight loss by assigning a dollar value to the externality and considering the pattern of substitution across vehicles. We begin with the 1988 through 1992 model years (as in Table 1 above), computing the possible welfare gains from a product-level tax and the deadweight loss from the second-best tax based on fuel economy. We then narrow the focus to demand in a single model year, using 1990 as an example, in order to explore the influence of a range of substitution patterns across vehicles. We show how certain correlations in the off-diagonal elements of the demand derivatives, $\partial x_j / \partial t_k$, can influence the fraction of welfare recovered in the second best.

Evaluating the formula in equation (6) first requires estimates for the externality generated by each car as well as its own- and cross-price elasticities with respect to the other vehicles in the market. We assign a value of \$39 for the social cost of carbon (Interagency Working Group on Social Cost of Carbon (2013)), leading to an external cost of 34.6 cents per gallon.[19] Using our data on

[18] We have run age-specific regressions in which we compute $R^2$ using data only for ages 10,...,17. $R^2$ decreases from 0.37 for 10-year-old vehicles to 0.22 for 17-year-old vehicles.

[19] If the cost associated with carbon emissions has been rising approximately at the discount rate, we interpret this

lifetime fuel use this implies an average of $3,172 in external costs for each vehicle sold. We further impose an own-price elasticity of -5 (roughly comparable to the estimates in Berry, Levinsohn, and Pakes (1995)) and cross-price elasticities distributed evenly over the full set of models. We relax both of these assumptions below, considering higher and lower own-price elasticities and cross-price elasticities that are correlated with similarity in attributes.

As above we compute welfare results relative to a baseline that controls for substitution to an outside good (since a revenue-neutral CAFE standard does not directly incentivize switching to an outside good) and so isolate the welfare effects coming from switching among vehicles. Under our assumptions on elasticities the welfare gains with a separate Pigouvian tax on each of the 1,636 VIN10-prefixes amount to $239 per car sold, or about $3.4 billion per year. The best linear tax on fuel use per mile, equivalent here to the optimal average fuel economy standard, generates about $0.75 billion per year in surplus and so leaves $2.6 billion in deadweight loss. This corresponds directly to the intuition on $R^2$ above: for the 1988-1992 model years the weighted $R^2$ is 0.22, implying 22% of possible gains are recovered with a single linear policy.

The single linear policy over the five model years also contains an inefficiency related to the differing sets of fuel economies available each model year. As an alternative, we repeat the calculation for a more flexible policy that updates each year. The fraction of first-best welfare recovered increases only slightly, to 23%.

Table 4 repeats the welfare calculation, now exploring sensitivity to the own- and cross-price elasticities across cars. We focus on substitution across cars in a single model year, 1990, for clarity. The first panel under the central case considers changes in the own-price elasticity of demand for individual models (-5 in the central case). More elastic demand allows a larger change in the composition of the fleet and so greater welfare gains are possible in the first best. The ratio of second- to first-best welfare gains remains fixed at 0.24, the value of $R^2$ for the 1990 model year.

The remaining panels in Table 4 explore correlation in cross-price elasticities related to attributes of the vehicles. Recall that the derivation of $R^2$ as a summary statistic involves own-price elasticities, with idiosyncratic variation in the cross-price effects canceling out. Strong correlation in cross-price effects (particularly with respect to the externality) can influence the pattern of welfare effects. Panel B allows cross-price effects to be related to the relative fuel economy of vehicles, with the two cases differing in how rapidly substitutability decays as cars become more distant. A car with the same fuel economy is twice as good a substitute as a car one standard deviation away (about 1 gallon per hundred miles) in the first case, and five times as good a substitute in the second case. When introducing this substitution pattern the gains possible in the first best become smaller. Intuitively, it is now harder to move people across vehicles with different externalities, at least to the extent the externality is correlated with fuel economy. This change in cross-price effects worsens the performance of the second-best policy even more dramatically than the first-best: this is because the only margin on which the second-best policy operates (fuel economy) is the margin that the cross-price effects are limiting. Panel C repeats the experiment, this time

value as being in 2011 dollars (looking retrospectively at the 1988-1992 vintages).

with cross-price elasticities such that cars with similar lifetime miles are the best substitutes. This pattern instead limits the importance of heterogeneity in durability since substitution across cars of different durabilities is slow even in the first best. The relative performance of the second-best policy is therefore improved. Finally, Panel D introduces substitution correlated with vehicle prices. Since price is not as strongly correlated with the externality as fuel economy or durability (which combine to define the externality) the effects on welfare in this case are quite small.

**Table 4:** Welfare Effects for Model Year 1990

| | **Second best** | **First best** | **Ratio** |
|---|---|---|---|
| Central case | 0.81 | 3.39 | 0.24 |
| A Own-price elasticity | | | |
| -4 | 0.66 | 2.73 | 0.24 |
| -6 | 0.97 | 4.03 | 0.24 |
| B Cross-price elasticities, fuel economy | | | |
| 2 | 0.71 | 3.30 | 0.21 |
| 5 | 0.55 | 3.15 | 0.17 |
| C Cross-price elasticities, lifetime miles | | | |
| 2 | 0.87 | 2.89 | 0.30 |
| 5 | 0.88 | 2.29 | 0.38 |
| D Cross-price elasticities, vehicle price | | | |
| 2 | 0.81 | 3.44 | 0.23 |
| 5 | 0.79 | 3.50 | 0.23 |

Note: Welfare gains are expressed in billions of 2011 dollars relative to a constant tax at the average externality. Cross-price elasticities are expressed as the factor by which substitutability decreases per standard deviation difference in the specified attribute.

The estimates above are subject to several important caveats. Perhaps the most important consideration is the role of technology: we consider a static portfolio of durables offered for sale while in the long run the products can be re-engineered according to incentives provided by the tax schedule. In the case of cars this amounts to altering attributes, for example reductions in weight and horsepower, and introducing efficiency-enhancing engine technologies. The second-best intuition developed here will also apply to the distribution of these technologies across cars: the most advanced technologies and lightest materials should be placed in the cars that have the highest expected lifetime mileage. A linear tax (or standard) based on fuel economy will encourage these technologies to be distributed much more equally across the fleet, missing welfare gains possible if the improvements could instead be targeted. Interactions between the second-best effects here and a set of other distortions produced by standards (for example due to attribute-basing or changes to durability induced via the used market as discussed in Jacobsen and van Benthem (2015)) also have the potential to influence welfare.

# 5 Model with heterogeneous consumers

In Section 2 we present a model in which consumers are all alike, but there is heterogeneity in lifetime utilization of durable goods due to differences in product durability. We now generalize to a setting in which consumers have different intensities of use of the same good and, therefore, each consumer-good pairing can have a different lifetime externality.

## 5.1 The Standard Diamond Model

Diamond (1973) derives the second-best tax $t^*$ for the case where there is a single externality-generating good $x$ but there are $i = 1, \dots, I$ consumers who have different own-price derivatives of demand and generate different externalities.

$$t^* = \frac{-\sum_{i=1}^{I} \sum_{k \neq i} \frac{\partial U^k}{\partial x_i} \frac{\partial x_i}{\partial t}(p+t)}{\sum_{i=1}^{I} \frac{\partial x_i}{\partial t}(p+t)} \tag{15}$$

The marginal damages (externality) is expressed as the impact of person $i$'s consumption on others $-\sum_{k \neq i} \frac{\partial U^k}{\partial x_i} \equiv e_i$. Equation (15) has a simple interpretation: the second-best tax is the average of the externalities, weighted by the individuals' own-price derivatives. If more responsive consumers have higher damages $e_i$, the second-best tax rate exceeds the unweighted average externality $\bar{e}$. If demand derivatives are uniform across consumers, $t^* = \bar{e}$. Note that if consumers are identical and have the same quasi-linear utility function ($U^i = U \ \forall i$) then the optimal demands $x_i^* = x^*$ and the externality $e_i = e$ for all consumers. Also, $\frac{\partial U^i}{\partial x_k} = \frac{\partial U^k}{\partial x_i} \ \forall i \forall k$ and $\frac{\partial x_i}{\partial t}(p+t) = x' \ \forall i$. In that case, the problem collapses to a first-best solution in which the optimal tax is equal to the total externality imposed on all other consumers: $t^* = e$.

## 5.2 A Diamond Model with Multiple Goods

When consumers have a different intensity of use of the same durable good, they will produce different externalities, $e_{ji}$, from an ex-ante identical product $j$. For example, different car owners may drive different amounts of miles over the lifetime of the same model.[20] Another example is that some drivers are more aggressive than others and as such realize a lower fuel economy and higher emissions (Langer and McRae (2014)). Another source of heterogeneity for local pollutants is geography; the same tailpipe emissions will have different external damages depending on population density and the baseline level of ambient pollution concentrations.

We now generalize the Diamond model to include multiple goods $j = 1, ..., J$. We continue to abstract from the mileage rebound effect, as our aim is to compare different types of product-based

[20] To the extent that vehicle depreciation is a function of mileage, rather than time, different annual VMT is not enough to imply heterogeneity. Consumer heterogeneity comes from different accident risk or time-based depreciation.

policies. The setup is quite general: every consumer can have unique expected damages when paired with each product.[21] This encompasses heterogeneity due to breakdowns or accidents.[22] Each consumer $i$ holds a continuous bundle of goods, where demand for good $j$ by consumer $i$ is denoted $x_{ji}$. In the context of vehicles, we can interpret consumer $i$ as a "driver type" that holds a portfolio of vehicles. There are $J$ second-best tax rates $t_j^*$, one for each good, whereas the first-best solution (which we assume to be infeasible) would be $J \times I$ tax rates $t_{ij}^* = e_{ji}$. Note that the latter set of taxes, which are unique to a consumer-product pairing, are equivalent to a gasoline tax as long as we abstract from the mileage rebound effect.

If a first-best (gasoline) tax were feasible, elasticities are irrelevant. The only information requirement is the marginal externality per unit of energy consumption. The first-best tax guarantees that all consumers pay exactly their total externality and choose their vehicles optimally (Pigou, 1932). In the second-best setting described here, elasticities do matter. The second-best taxes will be a function of the elasticity of substitution between vehicles, and the externality imposed by consumer $i$'s individual use of car $j$: different people drive the same models differently (i.e., they have different ex-ante VMT distributions for the same vehicle at the moment of purchase).

We now derive expressions for the second-best taxes with heterogeneous consumers and multiple goods. In the Diamond framework, utility is given by:[23]

$$U^i(x_{11}, ..., x_{ji}, ..., x_{JI}) + \mu_i, \tag{16}$$

where $\mu_i$ is income spent on other goods by consumer $i$. We further follow Diamond in assuming that externalities are diseconomies:

$$\frac{\partial U^i}{\partial x_{jk}} \leq 0, \tag{17}$$

for $k \neq i$. The marginal utility of consumption is independent of the demands of other consumers:

$$\frac{\partial^2 U^i}{\partial x_{ji} \partial x_{jk}} = 0, \tag{18}$$

for $k \neq i$ and for all goods $j$. Each consumer $i$ maximizes utility subject to a budget constraint:

[21] Restricted versions of this model could assume that each consumer carries a fixed amount of "utilization" (e.g., vehicle miles traveled), regardless of the product type (car) (s)he owns. In this case, the remaining individual-level heterogeneity comes from driving style and accidents. In general, the dimensions of the problem might be reducible depending on the shapes of the person-car externalities $e_{ji}$. If each consumer has a fixed distribution of lifetime mileage $VMT_i$ and each vehicle has a fixed fuel economy $MPG_j$, then $e_{ji} = VMT_i/MPG_j$ (ignoring discounting). We can capture this with fewer than $J \times I$ taxes.

[22] The marginal damage can be interpreted as relative to the risk-adjusted mean damages.

[23] Note that in this specification, consumers derive utility from owning cars. *Services* derived from owning cars (VMT) are not explicitly modeled. However, this framework allows for different drivers of the same car to impose a different externality on other consumers. Implicitly, we can think of this as heterogeneity in VMT. In fact, in this general framework, consumer $i$ driving car $j$ may impose a different externality on consumer $k$ vs. consumer $l$. This is relevant for local pollutants, but not for a uniformly mixed pollutant such as $CO_2$.

$$\max_{x_{1i},...,x_{Ji}} U^i(x_{11}, ..., x_{ji}, ..., x_{JI}) + \mu_i, \tag{19}$$

subject to:

$$\sum_{j=1}^{J}(p_j + t_j)x_{ji} + \mu_i = m_i, \tag{20}$$

where $m_i$ is total income. The first order conditions assuming an interior solution and using equation (18) are given by:

$$\frac{\partial U^i}{\partial x_{ji}}(x_{1i}, ..., x_{Ji}) = p_j + t_j = \tilde{p}_j, \tag{21}$$

for all goods $j$. The optimal choices for consumer $i$ depend not only on the tax-inclusive price of the durable good $\tilde{p}_j$, but also on the consumption level – and therefore the tax-inclusive price – of the other durable goods:

$$x^*_{ji} = x_{ji}(\tilde{p}_1, ..., \tilde{p}_J). \tag{22}$$

Analogous to Diamond (1973), we derive the optimal tax rate by writing down an expression for total utility:

$$W(t_1, ..., t_J) = \sum_{i=1}^{I} U^i(x_{11}(\tilde{p}_1, ..., \tilde{p}_J), ..., x_{ji}(\tilde{p}_1, ..., \tilde{p}_J), ..., x_{JI}(\tilde{p}_1, ..., \tilde{p}_J)) + \sum_{i=1}^{I} \mu_i. \tag{23}$$

Using equation (20) and assuming that tax revenues are returned lump-sum, we can rewrite this as:

$$\begin{aligned} W(t_1, ..., t_J) = &\sum_{i=1}^{I} U^i(x_{11}(\tilde{p}_1, ..., \tilde{p}_J), ..., x_{ji}(\tilde{p}_1, ..., \tilde{p}_J), ..., x_{JI}(\tilde{p}_1, ..., \tilde{p}_J)), \\ &+ \sum_{i=1}^{I} \left( m_i - \sum_{j=1}^{J} p_j x_{ji}(\tilde{p}_1, ..., \tilde{p}_J) \right). \end{aligned} \tag{24}$$

To find the second-best optimal vector of tax rates $(t^*_1, ..., t^*_J)$, we need to calculate first order conditions by differentiating equation (24) w.r.t. to each of the $t_j$'s:

$$\frac{dW}{dt_j} = \sum_{l=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{I} \frac{\partial U^i}{\partial x_{lk}} \frac{\partial x_{lk}}{\partial t_j}(\tilde{p}_1, ..., \tilde{p}_J) - \sum_{l=1}^{J}\sum_{i=1}^{I} p_l \frac{\partial x_{li}}{\partial t_j}(\tilde{p}_1, ..., \tilde{p}_J) = 0. \tag{25}$$

Now use the FOCs in equation (21) to rewrite this as:

$$\frac{dW}{dt_j} = \sum_{l=1}^{J} \sum_{k=1}^{I} \sum_{i \neq k} \frac{\partial U^i}{\partial x_{lk}} \frac{\partial x_{lk}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J) + \sum_{l=1}^{J} \sum_{i=1}^{I} t_l \frac{\partial x_{li}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J) = 0. \tag{26}$$

In sum, the formula equalizes the change in consumer surplus for both cars and the change in total externalities when $t_j$ is raised.

This is a fairly intuitive expression. The first term in equation (26) is the total net change in the externality from an increase in $t_j$. If cars $j$ and $l$ are substitutes, this is an improvement from car $j$ and a worsening from car $l$. The double summation over $k$ and $i$ summarizes the total externality imposed on all others for each consumer $k$. The marginal utility gains from reduced externalities reflect the benefits from raising the tax on car $j$. The second term is the reduction in consumer surplus for all cars and consumers and reflects the costs of raising $t_j$. In sum, we set $t_j$ to equalize the change in net consumer surplus for all cars ("costs") and the change in net externalities ("benefits").

For the 2-good case (we are still confirming the proper notation for the $J$-good case), solving the system of 2 equations yields the following expression:

$$\begin{aligned} t_1^* &= \left[ \sum_i \frac{\partial x_{2i}}{\partial t_1} \sum_k \sum_{i \neq k} \left( \frac{\partial U^i}{\partial x_{1k}} \frac{\partial x_{1k}}{\partial t_2} + \frac{\partial U^i}{\partial x_{2k}} \frac{\partial x_{2k}}{\partial t_2} \right) - \sum_i \frac{\partial x_{2i}}{\partial t_2} \sum_k \sum_{i \neq k} \left( \frac{\partial U^i}{\partial x_{1k}} \frac{\partial x_{1k}}{\partial t_1} + \frac{\partial U^i}{\partial x_{2k}} \frac{\partial x_{2k}}{\partial t_1} \right) \right] \frac{1}{\Delta}, \\ t_2^* &= \left[ \sum_i \frac{\partial x_{1i}}{\partial t_2} \sum_k \sum_{i \neq k} \left( \frac{\partial U^i}{\partial x_{1k}} \frac{\partial x_{1k}}{\partial t_1} + \frac{\partial U^i}{\partial x_{2k}} \frac{\partial x_{2k}}{\partial t_1} \right) - \sum_i \frac{\partial x_{1i}}{\partial t_1} \sum_k \sum_{i \neq k} \left( \frac{\partial U^i}{\partial x_{1k}} \frac{\partial x_{1k}}{\partial t_2} + \frac{\partial U^i}{\partial x_{2k}} \frac{\partial x_{2k}}{\partial t_2} \right) \right] \frac{1}{\Delta}, \end{aligned} \tag{27}$$

where the arguments $(\tilde{p}_1, ..., \tilde{p}_J)$ are omitted for notational simplicity and:

$$\Delta = \sum_i \frac{\partial x_{i1}}{\partial t_1} (\tilde{p}_1, \tilde{p}_2) \sum_i \frac{\partial x_{i2}}{\partial t_2} (\tilde{p}_1, \tilde{p}_2) - \sum_i \frac{\partial x_{i2}}{\partial t_1} (\tilde{p}_1, \tilde{p}_2) \sum_i \frac{\partial x_{i1}}{\partial t_2} (\tilde{p}_1, \tilde{p}_2). \tag{28}$$

In the single good, standard Diamond model in equation (15), the second-best tax is the average of the externalities, weighted by the individuals' own-price derivatives. In the multiple goods case, this simple intuition no longer holds. $t_1^*$ now consists of two terms. The second term still contains the weighted sum of the externalities, but now also the weighted average externalities generated by the change in consumption of the second car when the first car is taxed: the weights for the externalities from the second car are *cross*-price derivatives. Next, the entire second term is multiplied by a scale factor (the sum of the own-price elasticities of car 2 across the consumers). The first term of the expression for $t_1^*$ is similar to the second term, but it reflects changes in externalities when the second car is taxed and its overall scale factor is a sum of cross-price derivatives. Hence, if the own-price elasticities are large relative to the cross-price elasticities, the second term dominates and the expression converges back to the simple Diamond model. $\Delta$ is a scale factor.

We conjecture that this can be simplified using matrix notation. Label two objects from equation (27) for ease of notation. First, $\Delta D_{jl}$ is the change in demand for good $j$ when the price of good $l$ changes. If $J = 2$, this is a $2 \times 2$ matrix. Second, $\Delta E_j$ is the total *net* change in externalities from a change in the price of good $j$. For example, $\sum_i \frac{\partial \alpha_{2i}}{\partial t_1} = \Delta D_{21}$ and $\sum_k \sum_{i \neq k} \left( \frac{\partial U^i}{\partial \alpha_{1k}} \frac{\partial \alpha_{1k}}{\partial t_2} + \frac{\partial U^i}{\partial \alpha_{2k}} \frac{\partial \alpha_{2k}}{\partial t_2} \right) = \Delta E_2$. For $J = 2$, we can write the vector of optimal taxes as $T^* = (\Delta D^{'})^{-1} \Delta E$.

$\Delta D$ is the standard market-level Marshallian substitution matrix. To calculate $\Delta D$, we need to know aggregate cross-price elasticities, but not *who* changed demand. $\Delta E$ is the *net* change in the externality resulting from a change in the price of a vehicle. Its element $j$ is the sum of each person $k$'s change in externality from a price increase of car $j$, which we denote as $\Delta e_{jk} \equiv \sum_{i \neq k} \left( \frac{\partial U^i}{\partial \alpha_{1k}} \frac{\partial \alpha_{1k}}{\partial t_j} + \frac{\partial U^i}{\partial \alpha_{2k}} \frac{\partial \alpha_{2k}}{\partial t_j} \right)$. This term can be negative or positive. For example, when $J = 2$ we expect $\frac{\partial \alpha_{11}}{\partial t_1} < 0$ and $\frac{\partial \alpha_{12}}{\partial t_1} > 0$): raising the price of car 1 will lower demand for car 1 but raise demand for car 2. The net effect on the externality depends on relative magnitude of the elasticities and externalities.

To calculate $\Delta e_{kj}$ we need to know the *correlation* between marginal damages and the own- and cross-price elasticities. Absent a homogeneity assumption, we need to know this for each individual. We therefore conclude that the informational requirements to calculate the second-best taxes for the general case of heterogeneous consumers and multiple goods in equation (27) are large. We interpret this as a "negative result": the second-best tax rates require substantially more information than a first-best gasoline tax. Moreover, it may be hard to estimate these correlations without highly granular individual-level and product-level data. We are developing simulations to assess how much welfare is lost by resorting to product-based standards that ignore heterogeneity in usage intensity, and we are simultaneously investigating whether or not an analytical expression for deadweight loss can be derived for this general case.

The heterogeneous consumers, multiple goods model simplifies when additional restrictions are imposed. We consider three special cases (see Appendix B for the derivations). First, when cross-price derivatives are all zero, the second-best taxes collapse to the single-good Diamond formula in equation (15) for each durable good $j$. Second, when there is no correlation between damages and elasticities, the second-best tax rate for each good $j$ is the average externality over all consumers $\bar{e}_j$. Hence, the optimal tax formula collapses back to the durability-only case presented in Section 2. Third, consider a case in which consumers are homogeneous in that different individuals $i$ and $k$ drive car model $j$ in an identical way, though cars have different fuel economy and durability (VMT). In other words, the externality from driving vehicle $j$ is the same across each pair of consumers. The optimal tax for each vehicle reduces to its expected externality: $t_j^* = e_j$. In other words, the optimal tax in the durability only model in Section 2 is now first-best (again, maintaining the assumption of no rebound effect).

## 6 Conclusion

Product-based policies are a ubiquitous policy tool for regulating energy-related externalities. Heterogeneity in the lifetime utilization of durables causes product-based policies to be inefficient, as compared to emissions taxes (or, for some pollutants, equivalent fuel taxes). We show that, under intuitive conditions in a model with a representative consumer, the $R^2$ from a regression of lifetime emissions on energy efficiency ratings across products maps directly into welfare loss and acts as a sufficient statistic for the use of restricted policies that place a linear subsidy on products based on their energy efficiency rating, as compared to the efficient Pigouvian benchmark.

We use detailed data on vehicle scrappage, miles traveled and fuel economy to estimate welfare losses of using product-based policies versus gasoline taxes to mitigate greenhouse gas emissions in the automobile market. We estimate that CAFE-style policies recover only about one quarter of the welfare gains that could be achieved by a gasoline tax. We also derive second-best tax rates in a more general model that allows for usage heterogeneity from differences across consumers as well as differences across products. There, we demonstrate that second-best policy design requires a vast amount of information that is not likely to be available to policy makers. We plan to use simulation to assess how much welfare is lost by resorting to optimal product-based standards. An emissions tax would create superior welfare gains while simultaneously requiring much less information.

## References


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

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 64 (4):841–890.

Borenstein, Severin. Forthcoming. "A Microeconomic Framework for Evaluating Energy Efficiency Rebound and Some Implications." *Energy Journal* .

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.

Callaway, Duncan, Meredith Fowlie, and Gavin McCormick. 2015. "Location, Location, Location? What Drives Variation in the Marginal Benefits of Renewable Energy and Demand-side Efficiency?" Manuscript: UC Berkeley.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics* 1 (1):451–488. URL http://ideas.repec.org/a/anr/reveco/v1y2009p451-488.html.

Cullen, Joseph. 2013. "Measuring the Environmental Benefits of Wind Generated Electricity." *American Economic Journal: Economic Policy* 5 (4):107–33.

Davis, Lucas W. and Matthew E. Kahn. 2010. "International Trade in Used Vehicles: The Environmental Consequences of NAFTA." *American Economic Journal: Economic Policy* 2 (4):58–82. URL http://www.aeaweb.org/articles.php?doi=10.1257/pol.2.4.58.

Diamond, Peter A. 1973. "Consumption Externalities and Imperfect Corrective Pricing." *The Bell Journal of Economics* 4 (2):526–538.

Feng, Ye, Don Fullerton, and Li Gan. 2013. "Vehicle choices, miles driven, and pollution policies." *Journal of Regulatory Economics* 44 (1):4–29. URL http://ideas.repec.org/a/kap/regeco/v44y2013i1p4-29.html.

Fullerton, Don and Sarah E. West. 2002. "Can Taxes on Cars and Gasoline Mimic an Unavailable Tax on Emissions?" *Journal of Environmental Economics and Management* 42 (1):135–157.

———. 2010. "Tax and Subsidy Combinations for the Control of Car Pollution." *The B.E. Journal of Economic Analysis & Policy* 10 (1):1–33. URL http://ideas.repec.org/a/bpj/bejeap/v10y2010i1n8.html.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven. 2014. "Consumer Valuation of Fuel Costs and teh Effectiveness of Tax Policy: Evidence from the European Car Market." Manuscript: University of Leuven.

Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review Papers & Proceedings* 54 (3):58–76.

Hendren, Nathaniel. 2013. "The Policy Elasticity." Manuscript: Harvard University.

Holland, Stephen P., Jonathan E. Hughes, and Christopher R. Knittel. 2009. "Greenhouse Gas Reductions Under Low Carbon Fuel Standards?" *American Economic Journal: Economic Policy* 1 (1):106–146.

Holland, Stephen P., Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates. 2014. "Measuring the Spatial Heterogeneity in Environmental Externalities from Driving: A Comparison of Gasoline and Electric Vehicles." Manuscript: UNC Greensboro.

Interagency Working Group on Social Cost of Carbon. 2013. "Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis." Http://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/technical-update-social-cost-of-carbon-for-regulator-impact-analysis.pdf.

Jacobsen, Mark R. and Arthur A. van Benthem. 2015. "Vehicle Scrappage and Gasoline Policy." *American Economic Review* 105 (3):1312–1338.

Knittel, Christopher R. and Ryan Sandler. 2013. "The Welfare Impact of Indirect Pigouvian Taxation: Evidence from Transportation." Manuscript: MIT.

Langer, Ashley and Shaun McRae. 2014. "Step on It: Evidence on the Variation in On-Road Fuel Economy." Manuscript: University of Arizona.

Pigou, Arthur C. 1932. *The Economics of Welfare*. London: Macmillan and Co., 4th edition ed.

Sallee, James M. 2011. "The Taxation of Fuel Economy." *Tax Policy and The Economy* 25:1–37.

Sallee, James M., Sarah E. West, and Wei Fan. 2015. "Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations." Manuscript, University of Chicago.

# A Appendix: Additional Empirical Results

Table A.1 shows the sensitivity of $R^2$ to different treatments of observations with very high VMT-at-death. The first two rows indicate that dropping observations with VMT $\geq 1,000,000$ miles hardly affects $R^2$. Rows 3-6 indicate that, starting from the full sample, winsorizing at progressively lower VMT levels slightly increases $R^2$. For example, in the fourth row, any observation that has a reported odometer rating above 600,000 miles is recoded as having exactly 600,000 miles. Its gasoline consumption is recalculated assuming the new odometer reading, and the observation is then averaged along with all other observations from the same VIN10-prefix. The table reports OLS and WLS results, restricting the sample to model years 1988 to 1992 and to VIN10-prefixes with at least 200 observed retirements.

**Table A.1:** Regression $R^2$ Using Winsorized Data

| **VIN-pre averages, model years 1988-1992, models with** $N \geq 200$ | **OLS** | **WLS** |
|---|---|---|
| All odometer readings | .29 | .22 |
| Drop if odometer $\geq 1,000,000$ miles | .28 | .22 |
| Winsorize at 1,000,000 miles | .28 | .22 |
| Winsorize at 600,000 miles | .30 | .23 |
| Winsorize at 500,000 miles | .32 | .25 |
| Winsorize at 400,000 miles | .37 | .30 |

Note: Table shows $R^2$ from regressions of average lifetime gallons consumed on fuel consumption rating. The unit of observation is a VIN10-prefix.

# B Appendix: Special Cases of the Multiple Goods Diamond Model

This appendix contains the derivations for the three special cases of the heterogeneous consumer, multiple good model presented in Section 5. These special cases may aid in interpretation of the general result.

## B.1 No Cross-Price Effects in Demand

First, we demonstrate the our results collapse to the one-good result from Diamond (1973) when cars are neither substitutes nor complements for each other. Assume that:

$$\frac{\partial x_{ji}}{\partial t_l} = \frac{\partial x_{li}}{\partial t_j} = 0, \tag{B.1}$$

for all pairs of goods $j$ and $l$. Demand for one good does not depend on the price of the other good. Then, equation (22) simplifies to:

$$x^*_{ji} = x_{ji}(\tilde{p}_j). \tag{B.2}$$

In that case, the problem is separable in the different goods. Equation (26) reduces to:

$$\frac{dW}{dt_j} = \sum_{i=1}^{I} \sum_{k \neq i} \frac{\partial U^k}{\partial x_{ji}} \frac{\partial x_{ji}}{\partial t_j}(\tilde{p}_j) + t_j \sum_{i=1}^{I} \frac{\partial x_{ji}}{\partial t_j}(\tilde{p}_j). \tag{B.3}$$

The second-best taxes $t^*_j$ reduce to $J$ separate Diamond-style formulas of externalities weighted by own-price derivatives:

$$t^*_j = \frac{\sum_{i=1}^{I} \sum_{k \neq i} \frac{\partial U^k}{\partial x_{ji}} \frac{\partial x_{ji}}{\partial t_j}(\tilde{p}_j)}{\sum_{i} \frac{\partial x_{ji}}{\partial t_j}(\tilde{p}_j)}. \tag{B.4}$$

## B.2 No Correlation Between Elasticities and Externalities

Next we show what happens if the demand elasticities are uncorrelated with the externalities. Specifically, assume that:

$$\sum_{k=1}^{I} \sum_{i \neq k} \left( \frac{\partial U^i}{\partial \alpha_{lk}} \frac{\partial \alpha_{lk}}{\partial t_j} \right) = \sum_{k=1}^{I} \frac{\partial \alpha_{lk}}{\partial t_j} \sum_{k=1}^{I} \sum_{i \neq k} \frac{\partial U^i}{\partial \alpha_{lk}} / I. \tag{B.5}$$

If we substitute these terms into the second-best tax rates for the 2-good case in equation (27), the result is:

$$t_1^* = \sum_i \frac{\partial x_{2i}}{\partial t_1} \left( \sum_k \sum_{i \neq k} \frac{\partial U^i}{\partial x_{1k}} / I \sum_{k=1}^{I} \frac{\partial x_{1k}}{\partial t_2} + \sum_k \sum_{i \neq k} \frac{\partial U^i}{\partial x_{2k}} / I \sum_{k=1}^{I} \frac{\partial x_{2k}}{\partial t_2} \right) \frac{1}{\Delta}$$
$$- \sum_i \frac{\partial x_{2i}}{\partial t_2} \left( \sum_k \sum_{i \neq k} \frac{\partial U^i}{\partial x_{1k}} / I \sum_{k=1}^{I} \frac{\partial x_{1k}}{\partial t_1} + \sum_k \sum_{i \neq k} \frac{\partial U^i}{\partial x_{2k}} / I \sum_{k=1}^{I} \frac{\partial x_{2k}}{\partial t_1} \right) \frac{1}{\Delta}, \tag{B.6}$$
$$= - \sum_k \sum_{i \neq k} \frac{\partial U^i}{\partial \alpha_{k1}} / I = \bar{e_1}.$$

Analogously, $t_2^* = \bar{e_2}$. Hence, with no correlation between elasticities and externality, the second-best tax formula reduces to taxing each good at its average externality, which is the same as the second-best tax rate that we derive in the representative consumer case in section 2.

### B.3 Homogeneous Consumers

Another possible step would be to assume that consumer types are homogeneous in the following way: different individuals $i$ and $k$ drive car model $j$ in exactly the same way. In other words, cars may differ on characteristics such as fuel economy and durability (VMT), but there is no within-model heterogeneity of driving habits. As a consequence, the externality from driving vehicle $j$ is the same across consumers. This restriction can be expressed as follows:

$$\frac{\partial U^i}{\partial x_{jk}} = \frac{\partial U^i}{\partial x_{jl}} = u_j^{'},$$
$$\frac{\partial U^i}{\partial x_{jk}} = \frac{\partial U^k}{\partial x_{ji}} = u_j^{'}, \tag{B.7}$$

for $k \neq l \neq i$. The first condition imposes that if consumers $k$ and $l$ drive the same car $j$, they impose the same externality $e_j$ on consumer $i$. The second condition says that $k$'s externality on $i$ from driving car $j$ is the same as $i$'s externality on $k$ from driving the same car.

Substituting these constraints in the optimal tax FOCs in equation (26) yields:

$$\frac{dW}{dt_j} = \sum_{l=1}^{J} \sum_{k=1}^{I} \sum_{i \neq k} u_l^{'} \frac{\partial x_{lk}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J) + \sum_{l=1}^{J} \sum_{i=1}^{I} t_l \frac{\partial x_{li}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J),$$
$$= (I-1) \sum_{l=1}^{J} \sum_{i=1}^{I} u_l^{'} \frac{\partial x_{li}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J) + \sum_{l=1}^{J} \sum_{i=1}^{I} t_l \frac{\partial x_{li}}{\partial t_j} (\tilde{p}_1, ..., \tilde{p}_J) = 0, \tag{B.8}$$

which implies:

$$t_j^* = -(I-1) u_j^{'} = - \sum_{k \neq i} \frac{\partial U^k}{\partial x_{ji}} = e_j. \tag{B.9}$$

Under this assumption, product-based taxes can yield the first-best policy, where the optimal taxes are given by $t_j^* = e_j$. Here, the first-best tax equals the optimal tax in the durability only model in Section 2, and the first-best is obtainable because of the assumption that all people generate a common externality when using each type of product.