On the Use of Sufficient Statistics to Evaluate Externality Correcting Policies

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

August 21, 2015

Abstract

Pigouvian taxes fully correct for market failures due to externalities, but actual policies are commonly forced to deviate from the Pigouvian ideal by administrative or political constraints. This paper develops a framework for deriving sufficient statistics, which require a minimum of market information, that aid in the evaluation of the efficiency costs of such constraints on policy design. We demonstrate that, under certain intuitive conditions, standard output from a regression of true externalities on policy variables, including the R^2 and the sum of squared residuals, have immediate welfare interpretations—they are sufficient statistics that compare alternative policies. We utilize our framework in three empirical applications that address diverse factors that cause actual policy to deviate from the Pigouvian ideal: random mismeasurement in externalities, imperfect spatial differentiation, and heterogeneity in the longevity of durable goods. The latter of these raises a set of concerns that is entirely new to the literature on energy efficiency policy—the externalities attending energy-consuming durable goods depend on both energy efficiency and total lifetime utilization, but policy is generally based on only energy efficiency. Using our framework and a novel data set, we find that policies that regulate vehicle fuel economy, but ignore the variation in average longevity across different types of automobiles, recover only about one-quarter to one-third of the welfare gains achievable by a policy that also takes product longevity into account.

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

---

*We thank Richard Blundell, Meghan Busse, Don Fullerton, Jeff Grogger, Ravi Kanbur, Henrik Kleven, Johannes Spinnewijn and seminar participants at the 2015 AEA meetings, the 2015 AERE conference, the Spring 2015 NBER EEE meetings, the 2015 Northeast Workshop on Energy Policy and Environmental Economics, Imperial College, the Institute for Fiscal Studies, LSE, UCL, UC Berkeley and the University of Chicago for helpful comments and suggestions.

[1]Department of Economics, University of California San Diego. [2]Sloan School of Management, Massachusetts Institute of Technology. [3]Department of Agricultural and Resource Economics, University of California, Berkeley. [4]The Wharton School, University of Pennsylvania. [5]National Bureau of Economic Research.

1 Introduction

Many important policies aim to fix market failures due to the existence of an externality. Examples range from taxes on cigarettes, alcohol or sugary beverages to mandatory immunizations to the regulation of pollution. Since Pigou (1932), economists have understood that where externalities can be taxed directly, market efficiency can be fully restored. Yet, relatively few policies follow that prescription closely. Often it is administratively impossible, or politically infeasible, to truly price actions according to the externalities that they create. Consequently, externality correcting policies are generally imperfect, and a common goal of economic research is to estimate the efficiency costs of these policy imperfections. The goal of this paper is to facilitate such exploration.

To do so, we develop a theoretical framework that identifies sufficient statistics that can be used to assess the welfare costs of externality-correcting policies that deviate from the Pigouvian benchmark. Our theoretical derivations reveal that, when certain conditions are met, familiar statistics from simple regressions of the true externality related to a product or action on the variables upon which policy is based have welfare interpretations. In particular, deadweight loss scales directly with the sum of squared residuals, and the R^2 summarizes the percentage of the first-best welfare gain that is achieved by a second-best policy. We demonstrate the use of our framework via three distinct empirical applications.

The imperfect targeting of an externality correcting policy can frequently be conceived of as a situation in which the externality attending a product or action is a function of a collection of variables, but policy is made to be contingent upon only a subset of those variables, or upon their imperfect proxies. For example, one goal of a tax on alcohol is to prevent fatalities from drunk driving. The risk associated with a particular alcoholic beverage depends on many factors, including the location of its consumption, the time of day, the weather, etc., but taxes are generally a function only of a product's alcohol content. Our goal is to assess the efficiency costs of policies that are limited to taxing alcohol content, that are unable to take weather, time of day, etc., into account. Given sufficient information about the market, the efficiency of any policy alternative can be estimated. Our framework sets out to understand when policy alternatives can be evaluated with only limited market information, in the spirit of Chetty (2009). We specify a standard model of a competitive market with a representative consumer who chooses among a variety of related goods that may each cause varying levels of an externality. A Pigouvian tax can achieve the first best allocation, but we suppose that the planner faces a constraint on the set of taxes that can be set. We follow Harberger (1964) in deriving a general equilibrium expression that characterizes the deadweight loss of some alternative set of taxes that deviates from the Pigouvian benchmark using a local approximation. Evaluating this full expression requires information about all cross-product demand derivatives, which will typically be unavailable. Our main theoretical contribution is in deriving different conditions under which this general equilibrium expression simplifies so that only a few pieces of information are needed for policy comparisons. All that is required is some information about the variation in the true externality and some estimate of the degree to which the variables that enter the policy function are correlated with the true externality. We demonstrate that, when key assumptions are met, standard output from linear regressions of the true externality on the characteristics in the policy function have welfare interpretations.

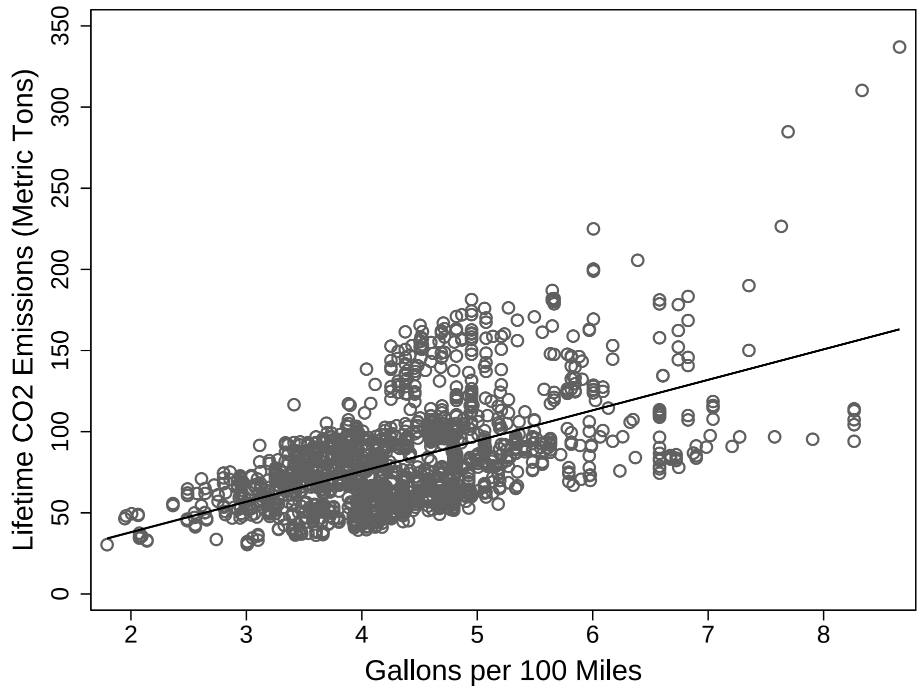
Whether the requisite assumptions are likely to hold in any given circumstance depends on market factors and the nature of the policy in question. To illustrate that our conditions may often be met, we use our framework in three empirical exercises. All three examples are drawn from environmental policy, but they span a variety of different situations (and reach different conclusions about the efficiency of constrained policies). One application considers random mismeasurement—energy efficiency is measured according to laboratory test procedures which differ from in-use averages, thereby creating mismeasurement across regulated products. We take advantage of a change in the fuel economy test procedure for automobiles in the United States to quantify the efficiency cost of basing fuel economy regulation on the older, noisier test procedure.

A second application considers spatial differentiation—a given amount of pollution may have quite different health or environmental consequences depending on the location at which it is emitted, but policies generally lack the ability to fully differentiate policy treatment of emissions across space. We use our framework to derive sufficient statistics that characterize the welfare costs of imperfect spatial differentiation and quantify those costs for the case of carbon dioxide emissions resulting from the use of home appliances, where differences in emissions across space are due to the fact that the marginal source of electricity has a different emissions rate in different parts of the country.

The third application concerns the regulation of energy-consuming durable goods that have heterogeneous total lifetime utilization. The lifetime pollution stemming from a durable good depends upon both its energy efficiency and its lifetime utilization, but policies that regulate energy efficiency ignore differences in product longevity. We use a novel data set that indicates the lifetime miles traveled for a large sample of automobiles to quantify the variance in average lifetime utilization of different types of automobiles. We find that average lifetime miles traveled for the individual vehicles of a particular model varies greatly across different models. This implies that vehicle models with the same fuel economy rating in fact have very different levels of expected lifetime carbon dioxide emissions. Fuel economy regulations, however, are forced to treat such vehicles identically.

We use our sufficient statistics approach and these data to conclude that policies, like fuel economy regulations, that ignore longevity differences and impose (implicit) taxes on automobiles according to only fuel economy recover only about one-quarter to one-third of the welfare gain from carbon mitigation achievable by a policy that considers both fuel economy and vehicle longevity. Our welfare conclusions come directly from running a bivariate linear regression of average lifetime emissions across models on official fuel economy ratings. This is illustrated in Figure 1. Each data point represents the average lifetime CO_2 emissions across a number of individual vehicles of the same model (e.g., a 2012 Toyota Camry). The solid line is the OLS regression line. The dispersion in the figure is driven by the substantial heterogeneity across models in average lifetime miles driven, which implies that policies based on fuel economy alone will be imprecise. The R^2

Figure 1: The Relationship Between Lifetime CO_2 Emissions and Fuel-Efficiency



Note: An observation is the average lifetime miles driven of a particular vehicle model, across many individual units, divided by the fuel economy rating and multiplied by the tons of CO_2 per gallon of gasoline. The sample is restricted to models for which we observe at least 200 retirements from model years 1988 to 1992. The data are described in detail in Section 3. The solid line is an OLS prediction line.

from this regression – 0.30 in the case of Figure 1 – is an estimate of the fraction of the first-best welfare gain that is achieved by a second-best fuel-economy policy.

There is a very large literature that studies the efficiency properties of energy efficiency policies, but, to the best of our knowledge, no prior paper has pointed to the welfare implications of heterogeneity in product longevity, which we find to be substantial. Related regulations that will suffer from the same inefficiency are central to energy policy. 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. And, a significant portion of energy policies are aimed at regulating the efficiency of these durables (rather than taxing fuel or emissions directly). Examples include fuel economy regulation, efficiency standards for appliances, 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] Our empirical examination of automobile longevity establishes a new factor that influ-

[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

ences the welfare properties of such policies, and, at least for the case of automobiles, we find that this factor is economically important.

Our paper contributes to the sufficient statistics tradition, which stretches at least to Harberger (1964) and was recently reviewed in Chetty (2009), in developing a framework for analyzing externalities and in deriving the relationship between simple regression statistics and welfare. One goal of our paper is to broaden the use of sufficient statistics for evaluating externality-mitigating policies, in particular in the areas of energy and the environment. While some theoretical papers in the energy and environmental literature can be thought of as taking a sufficient statistics approach, the connection between these literatures seems weak. Chetty (2009), for example, cites no papers focused on externalities, energy or the environment. A vast literature in energy and the environment considers second-best policy design, but it generally does so using specific models, structural assumptions and simulations in order to assess welfare.[2] That approach has the advantage of allowing for an internally consistent evaluation of a rich set of policy alternatives. Our approach differs in focusing on more general insights that can be gleaned with a minimum of modeling assumptions.

Our empirical applications also make specific contributions to the literature. Our study of automobile longevity adds a new insight to the substantial literature on energy efficiency policy (and fuel economy policy specifically) in showing how heterogeneity in expected product lifetime determines the welfare implications of policy.[3] Our study of spatial heterogeneity contributes to a growing empirical literature on the topic, none of which adopts a sufficient statistics approach.[4] A key theoretical contribution on spatial heterogeneity is Mendelsohn (1986), which pursues sufficient statistics (though it does not use the term) for evaluating imperfect environmental regulation. Relative to Mendelsohn (1986), we establish a more general framework that encompasses spatial heterogeneity as one example, and we derive specific results that relate welfare measures to regression statistics.

The balance of the paper is as follows. In Section 2 we develop our theoretical framework and derive sufficient statistics related to the correlation between externalities and product demand. In Section 3 we apply these results to the case of heterogeneity in the longevity of automobiles. Section 4 considers the case of spatial heterogeneity in emissions from identical products used in different locations. In this section, we derive a different set of sufficient statistics that are more

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 (2015) 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.

[2]Key examples in the literature related to automobiles are Fullerton and West (2010)

[3]Existing research, including Fullerton and West (2010), Fullerton and West (2002) and Feng, Fullerton, and Gan (2013), has considered how heterogeneity across consumers in driving behavior influences optimal policy design and welfare consequences, and Knittel and Sandler (2013) examine similar questions related to heterogeneity across individual automobiles in their local air pollution emissions rates, but we are aware of no prior studies considering how heterogeneity in average product durability affects policy and welfare.

[4]Holland, Mansur, Muller, and Yates (2014) studies the difference in emissions savings for subsidies for electric vehicles based on geography. Cullen (2013) and Callaway, Fowlie, and McCormick (2015) explore heterogeneity in emissions reductions in electricity generation from renewables or demand side management. Muller and Mendelsohn (2009), Muller, Mendelsohn, and Nordhaus (2011) and Fowlie and Muller (2013) all study heterogeneity in damages from air pollutants based on their location of emission.

likely to apply to cases of spatial heterogeneity, and we quantify them for the case of refrigerators that are ex ante identical but emit different levels of carbon dioxide when consumed in different parts of the country. In Section 5 we apply our framework to the case of random mismeasurement in externalities, using a recent change in fuel economy testing procedures for automobiles. Section 6 concludes.

2 Theory for Deriving Sufficient Statistics

The goal of our model is to facilitate analysis of the efficiency costs of externality-correcting policies that deviate from the theoretical ideal. Real-world policies may be less efficient than an ideal policy for a variety of reasons, including political constraints and administrative feasibility. Often policy limitations come in the form of dimensional simplicity: externalities are determined by several factors, but policy is contingent upon only a subset of them. For example, the external damages caused by air pollution depend upon the quantity of emissions, weather (which determines how pollutants travel), population density in surrounding areas, and baseline levels of various chemicals in the ambient environment. Policies that target air pollution—like the SO_2 trading program in the United States—tax industrial products in proportion to their emissions quantities only, ignoring the other factors. Making such a policy contingent upon all factors that determine damages may often be infeasible. Our aim is to establish conditions under which the efficiency costs of such a policy constraint can be understood with a minimum of market information; that is, we seek sufficient statistics for policy evaluation.

To develop our model, we first present our setup and notation. We then derive a general expression that approximates the deadweight loss of using some alternative policy in lieu of the ideal. We then specify assumptions under which this general expression collapses to a more simple result that requires only a few pieces of information about the market.

To better isolate our phenomenon of interest, we emphasize a simple model in which the only market failure is the externality to be targeted—we assume perfect competition, a representative consumer and no technological change. In this setting, a Pigouvian tax on the products in the market will achieve the first-best allocation, and we evaluate alternative policies against this theoretical benchmark. In choosing which alternative policies to evaluate, we focus on second-best policies that are chosen against some exogenous constraint. We describe policies as taxes on products, but this is equivalent to regulatory policies that create implicit taxes (shadow prices) on products.

2.1 Model setup

We model a representative consumer in a perfectly competitive market. The economy has products indexed $j = 1, ..., J$. The consumer chooses quantities of each, denoted x_j. Consumers derive utility, U, from the consumption of these products according to the function $U(x_1, ..., x_J)$, which we assume is twice differentiable, increasing and weakly concave in each argument. We denote the cost of production by $C(x_1, ..., x_J)$, which we assume is twice differentiable, increasing and weakly convex in each argument. There is an exogenous amount of income in the economy, M, and all remaining income is consumed in a quasi-linear numeraire, n. We do not model the endogenous entry and exit of products into the market—as such, ours is a short-run model—though it is not difficult to modify our model to allow for zero quantities so that the product vector represents potential products.[5]

Each product may contribute to a social externality, denoted ϕ. We assume the externality is a linear function of the total consumed of each good. That is, $\phi = \sum_{j=1}^{J} \phi_j x_j$, where ϕ_j is the marginal damage per unit of good j.[6] We assume that ϕ_j is fixed—that is, the marginal externality from each product is not itself a function of the policy regime. This is a natural assumption in many situations, but it has some particular implications for comparing first and second-best policies when the goods in question are durable goods that consumers can choose to use more or less intensively. We return to these issues where relevant in our empirical applications, and we enumerate this assumption for ease of later reference.

Assumption 1. *Marginal social damages from each product, ϕ_j, are fixed.*

A natural way to think of our setup is that it models a sector of the economy—e.g., j indexes types of refrigerators, and n is a separable bundle that represents all other goods. Each of the goods in the sector contributes varying amounts, ϕ_j, to a common externality—e.g., the use of each refrigerator over its lifetime leads to a different amount of carbon dioxide, discounted to the present. The consumer ignores the externality when making choices, and the goal of the planner is to use taxes to internalize the externality.

The planner can impose product taxes, denoted t_j. With competitive supply, the equilibrium price of a product to suppliers, denoted p_j, will equal marginal cost. We assume that consumers remit taxes, so that the price to consumers is $p_j + t_j$. Revenue is recycled lump-sum to consumers through a grant D. The consumer acts as a price taker and does not account for the externality when making her choice. The consumer's optimization problem is:

$$\max_{x_1,...,x_J} Z = U(x_1, ..., x_J) + n \quad \text{s.t.} \quad \sum_{j=1}^{J} (p_j + t_j)x_j + n \le M + D. \tag{1}$$

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.

Social welfare W is the utility from the product bundle, the numeraire (substituted out for the

[5]For notational ease, we assume non-zero quantities throughout, so that all first-order conditions are equalities.

[6]Note that for externalities with nonlinear effects, this setup can be understood as a local approximation.

budget constraint), and the externality:

$$W = U(x_1, ..., x_J) + M - C(x_1, ..., x_J) - \sum_{j=1}^{J} \phi_j x_j. \tag{2}$$

We say the planner is unconstrained when she can set a unique tax rate on each product. In this case, the planner's problem is:

$$\max_{t_1,...,t_J} W = U(x_1, ..., x_J) + M - C(x_1, ..., x_J) - \sum_{j=1}^{J} \phi_j x_j. \tag{3}$$

The first-order condition for product j is:

$$\frac{dW}{dt_j} = \sum_{k=1}^{J} \left(\frac{\partial U}{\partial x_k} - \frac{\partial C}{\partial x_k} - \phi_k \right) \frac{\partial x_k}{\partial t_j} = \sum_{k=1}^{J} (t_k - \phi_k) \frac{\partial x_k}{\partial t_j} = 0, \tag{4}$$

where the second equality follows from substituting in from the consumer's first-order condition.

All J first-order conditions for the planner will be met if and only if $t_j = \phi_j \ \forall j$. That is, the planner's solution is a vector of Pigouvian taxes; each product's tax rate is set equal to its marginal external damage. This is to be expected, as our setup is a standard tax model. The Pigouvian tax vector is our benchmark policy. We next derive an expression for the deadweight loss of imposing some alternative tax schedule in lieu of this benchmark.

2.2 Characterizing deadweight loss

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. Let any generic tax schedule be denoted as $\tau_1, ..., \tau_J$. We characterize the welfare loss of moving from the ideal 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 ϕ_j to τ_j. We denote the difference in welfare between the two schedules as $DWL(\tau)$.

In line with the sufficient statistics literature, we assume that demand derivatives are constant over the relevant range of taxes. Where demand derivatives are non-constant, this can be interpreted as a local approximation.

Assumption 2. *Demand derivatives* $\frac{\partial x_j}{\partial t_k}$ *are constant between* ϕ_j *and* τ_j *for all* j *and* k.

Under the assumption of constant demand derivatives, the efficiency loss incurred from imposing any arbitrary tax schedule τ in lieu of the first-best Pigouvian tax schedule can be written as:

$$W(t = \phi) - W(t = \tau) \equiv DWL(\tau) = -\frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{J} (\tau_j - \phi_j) (\tau_k - \phi_k) \frac{\partial x_j}{\partial t_k}. \tag{5}$$

The proof, along with all others, is in Appendix A. This formula is in the form of a set of Harberger triangles, and indeed the same result (though without externalities) is in Harberger (1964). When $\tau_j = \phi_j$, each term in the summation will be zero. Note that the Harberger triangles include a complete set of both own-price effects and cross-price effects.

To better understand the content of equation (5) we substitute $e_j \equiv (\tau_j - \phi_j)$, where e_j is the "error" in the tax rate, and decompose the own and cross terms:

$$-2 \times DWL(\tau) = \sum_{j=1}^{J}\sum_{k=1}^{J} e_j e_k \frac{\partial x_j}{\partial t_k} \tag{6}$$

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

Formulas (6) and (7) are quite general. But, using these formulas to evaluation policy alternatives requires knowledge of the complete demand matrix, including all cross-price derivatives. This information will frequently be unavailable.

Under some conditions, however, the expression will simplify further and policy evaluation will require less information. In particular, in some circumstances (which we detail below), the cross-effects in equation (7) will cancel to zero. Then, deadweight loss will depend only on the squared tax "errors" and own-price derivatives (which are negative and thus sum to a positive deadweight loss). To see that this is possible, suppose that the vector τ is "unbiased", so that, on average across j, the tax rate is equal to marginal damages. Cross-effects are then the sum over a mean zero term, and under an assumption about the correlation between tax "errors" and the substitution matrix, these effects will sum to zero. We next specify conditions under which this cancellation occurs. Later, when considering our second empirical application in Section 4, we derive a different set of sufficient statistics for a case when this cancellation does not occur, but a different assumption about discrete choice models leads to an alternative simplification.

2.3 Sufficient statistics for second-best policies

Equation (6) characterizes the welfare loss incurred from having a tax vector that differs from the Pigouvian benchmark. Our interest is in policies that arise from some constraint on the tax vector, and we focus on policies that are second-best given some constraint. Constraints on the tax policy can be described quite generally as follows. Denote some vector of characteristics $\mathcal{A}$ that determine a product's externality: $\phi_j(\mathcal{A})$. We suppose that the tax rate is a function of some other (possibly overlapping) vector of attributes $\mathcal{B}$: $\tau(\mathcal{B})$. For example, the lifetime carbon emissions from an automobile depend upon both its fuel economy rating and the number of miles it is driven over its life, but fuel economy policies are based only on fuel economy ratings. In this case, $\mathcal{A} = \{$fuel economy, lifetime mileage$\}$, and $\mathcal{B} = \{$fuel economy$\}$. Again, we suppose that $\mathcal{A}$ and $\mathcal{B}$ may differ because of administrative feasibility, political constraints, or other factors. We take as given that there is some reason for imperfect policy design.

For expositional ease, we derive results for the case where $\mathcal{B}$ is a single exogenous variable, denoted f_j, and the tax policy takes the form of a linear function of f_j. Then the policy choice is to choose α and β where $\tau_j = \alpha + \beta f_j$. It is straightforward to modify our derivation to include many variables, but this comes at the cost of considerable notational complexity. In terms of the linear form of the tax, what is critical is that the tax policy can be written as linear in parameters for any combination of the set of variables included in $\mathcal{B}$—including higher-order transformations and interactions. The linear in parameters assumption is important because we relate deadweight loss to ordinary least squares.

The planner's second-best problem is to choose α and β to minimize deadweight loss:

$$\min_{\alpha,\beta} DWL(\tau) = -\frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}(\phi_j - \alpha - \beta f_j)(\phi_k - \alpha - \beta f_k)\frac{\partial x_j}{\partial t_k}. \tag{8}$$

One candidate solution to this problem will be to choose α and β via an ordinary least squares fit of ϕ_j to f_j. We define the "structural tax error" ε_j as the error for product j from the OLS fit of ϕ_j on f_j: $\varepsilon_j \equiv \phi_j - \alpha^{OLS} - \beta^{OLS} f_j$. We make use of this additional concept because ε_j is defined in terms of primitives—ϕ_j and f_j are exogenous to policy—whereas the errors in tax rates e_j are endogenous to policy. Note, however, that ε_j depends on the variable f_j—the structural error represents the portion of the externality that is not predicted by the variable upon which policy is based. In the example cited above, f_j would be fuel economy, and ε_j would represent the portion of lifetime carbon emissions for vehicle j that is not predicted by its fuel economy, which depends upon how vehicle j's lifetime mileage differs from the average.

When the structural errors are uncorrelated with product substitutability, the second-best policy will be the solution to a linear best fit of the externality to the attribute, and when the second-best policy is implemented, the cross effects in equation (7) will cancel. Intuitively, the deadweight loss of the imperfect second best tax rates in equation (8) are scaled by the size of the own-price derivatives, which is a standard result in public finance. The solution simplifies still further when the policy attribute f_j is uncorrelated with own-price derivatives. We state this formally in assumptions 3 and 4:

Assumption 3. *Structural errors are uncorrelated with own-price derivatives: $cov(\varepsilon_j, \frac{\partial x_j}{\partial t_j}) = 0$; and, for each j, cross-price derivatives are uncorrelated with the structural errors of its alternatives: $\sum_{k\neq j}^{J} \frac{\partial x_j}{\partial t_k}\varepsilon_k = 0 \ \forall j$.*

Assumption 4. *The own-price derivatives for products are uncorrelated with f_j: $cov(f_j, \frac{\partial x_j}{\partial t_j}) = 0$.*

Assumption 4 is useful for expositional clarity, but it is straightforward to relax (we do so below), while assumption 3 is worthy of greater discussion. The first part of assumption 3 just says that we assume that the strength of own-price derivatives is not correlated with a product's structural error; that is, whatever factors are omitted from the policy function that determine the externality do not also indicate stronger or weaker own-price responses.

The second part of assumption 3 has more economic content; it says that, for a given product j, among that product's substitutes, there is no correlation between tax errors and their relative substitutability. That is, from the perspective of demand for product j, the tax errors are randomly distributed across its stronger and weaker substitutes. This part of the assumption is actually stronger than necessary—what is required is for this condition to hold on average across products, not for every j—but we also require two variants of this condition to also be true, all of which are implied by the version stated in assumption 3.

Assumption 3 will be met when the difference between the structural errors in the tax rates between two products j and k is no smaller or larger when the two products are closer substitutes. The errors in tax rates represent the residual variation in the externality, after conditioning on attribute upon which policy is contingent, f. Consider the vehicle example. Two vehicles with similar externalities (ϕ) will be closer substitutes, provided that vehicle fuel economy (f) is a factor that determines vehicle choice, because ϕ is mechanically related to f. But, assumption 3 can still be met if, after conditioning on fuel economy, the residual variation in ϕ is not correlated with substitutability.

Put differently, assumption 3 will be met if, from the perspective of consumer demand, conditional on the variables used to determine policy, differences in the externality across products are random. Whether this will be true depends upon the variables that are included in the policy and the source of residual variation in the externality. The validity of assumption 3 is an empirical question, which – as we will argue in Sections 3 and 5 – is likely to be a reasonable approximation for fuel-economy standards for vehicles or policies based on noisy energy efficiency ratings. For other applications, such as taxes on electricity-consuming home appliances, the cross-effects will not be zero but we can still derive a different sufficient statistic by imposing realistic regional structure on the demand matrix (Section 4).

When assumptions 1 to 4 hold, the second-best policy (the solution to the problem stated in equation (8)) will be to choose α and β to be the OLS solutions from fitting the externality to the policy variable. This is stated in Proposition 1:

Proposition 1. *Under assumptions 1 to 4, the second-best policy is the OLS fit of ϕ_j to f_j, and the deadweight loss is equal to the sum of squared residuals, multiplied by the average own-price derivative:*

$$DWL = -\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}\sum_{j=1}^{J}(\phi_j - \alpha^{OLS} - \beta^{OLS} f_j)^2 \tag{9}$$

$$= -\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}} SSR. \tag{10}$$

The proof is in Appendix A. The intuition is as follows. When the externality, conditional on characteristics that are in the policy function, are uncorrelated with product substitutability, then cross-effects will be zero on average. The partial equilibrium Harberger triangle is equal to the general equilibrium Harberger triangle that recognizes cross-market effects. The second-best policy is thus to minimize the sum of the own-price Harberger triangles, which is done when the policy employs a best-fit approach. A simple OLS fit is optimal when the own-price demand derivatives are uncorrelated with the policy variable.

In turn, the resulting errors in tax rates are the residuals from an OLS regression, and the deadweight loss is the sum of squared residuals from the OLS regression, scaled by the average demand derivative. With information about the average own-derivative and data on actual externalities and the attribute upon which policy is based, one can estimate deadweight loss of the policy constraint. The average own-price derivative of products and the SSR are sufficient statistics.

Moreover, the R^2 from this regression is a sufficient statistic that summarizes the percentage of welfare gain that could be achieved by the Pigouvian benchmark that is achievable by the second-best constrained policy. The percentage gain in welfare must be defined relative to some benchmark. The R^2 is defined relative to a benchmark policy that imposes a single unbiased tax rate on all products.

Corollary 1. *Under assumptions 1 to 4, the R^2 from the OLS fit of ϕ_j to f_j represents the percentage of the welfare gain of the Pigouvian tax (relative to a baseline of a second-best constrained to tax all products equally) that is achieved by the second-best linear tax on f_j (relative to the same baseline):*

$$R^2 = \frac{DWL(\tau = \alpha^{OLS} + \beta^{OLS} f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)}. \tag{11}$$

Note that the R^2 relaxes the information requirement of knowing the own-price derivatives. No moments of the demand system are required to calculate this sufficient statistic (though it hinges on the veracity of assumptions about the demand system). This makes assessing the relative welfare gain very intuitive and easy: all that is required is running a simple OLS regression of the actual externality for each product on the variables used in the policy function!

A version of our result that relaxes assumption 4 also has an intuitive relationship to standard regression analysis.

Proposition 2. *Under assumptions 1 to 3, the second-best policy is the generalized least squares fit of ϕ_j to f_j, where the weighting matrix is diagonal with each entry equal to the own-price derivative for product j. The deadweight loss is the elasticity-weighted sum of squared residuals from the GLS estimator:*

$$DWL = -\frac{1}{2} \sum_{j=1}^{J} (\phi_j - \alpha^{GLS} - \beta^{GLS} f_j)^2 \frac{\partial x_j}{\partial t_j}. \tag{12}$$

Given information about the own-price derivatives of each product, a researcher could calculate the GLS estimator and derive parallel welfare results for this case, which relaxes our assumption that own-price derivatives are uncorrelated with f_j. Further, when relaxing assumption 3, the second-best policy is the GLS fit of ϕ to f_j where the weighting matrix is the full demand matrix. We do not emphasize this result because it requires additional information about the demand system, but in many instances use of this formula would be useful for robustness analysis.

These main results demonstrate that—under some conditions about the demand system, most importantly that, conditional on the variables in the policy function, the externality is not correlated across products that are more substitutable—the deadweight loss of deviating from the Pigouvian benchmark can be calculated with limited information about the market. In the next three sections we demonstrate that this theoretical result has empirical relevance by illustrating three situations in which a sufficient statistic useful for evaluating policy can be derived from this framework.

3 Application 1: Automobiles and Longevity

Our first empirical application concerns the regulation of energy-consuming durable goods according to their energy efficiency ratings. Many of the most important energy policies—including fuel economy regulations, energy efficiency standards for appliances, and building codes—take this form. As discussed above in the introduction, policies in this genre have a number of known weaknesses; nevertheless, they remain ubiquitous. In this section, we point out a concern with such policies that is new to the literature—where the durables being regulated vary in their ex ante expected lifetime utilization, regulations based only on energy efficiency can lead to significant relative mis-pricing of products. We use a novel data set on the lifetime mileage of automobiles to calculate sufficient statistics that compare the welfare effects of a policy that takes both fuel economy and product longevity into account, as compared to one that considers only fuel economy. We focus on the implications of policy design for the mitigation of greenhouse gases from automobiles. We find large welfare implications—ignoring longevity leads to welfare costs on the order of $2 billion per year.

Specifically, we consider the welfare cost of a second-best tax on vehicles that is a linear function of fuel consumption (the inverse of fuel economy), whereas actual greenhouse gas externalities depend on both fuel consumption and lifetime miles driven.[7] In the notation of Section 2, $\mathcal{A} = \{$fuel consumption, lifetime mileage$\}$; $\mathcal{B} = \{$fuel economy$\}$; $\phi_j = \kappa f_j m_j$; and $\tau_j = \alpha + \beta f_j$, where f_j is fuel consumption in gallons per mile, m_j is lifetime miles driven, κ is the social cost of carbon multiplied by the tons of carbon per gallon of gasoline, and α and β are policy parameters to be chosen.

We use data from the California Smog Check program, merged with a national registration database that identifies when a vehicle has been retired from the U.S. fleet, in order to identify the lifetime miles traveled m_j for different types of automobiles. Note that we are concerned with heterogeneity in lifetime mileage at the vehicle type level (e.g., a 1995 Toyota Camry), not the

[7] Gasoline consumption is linearly related to fuel consumption, not fuel economy. Fuel economy regulations actually regulate average fuel consumption rates, not fuel economy rates, despite their name.

individual vehicle level (e.g., the 1995 Toyota Camry with serial number 123456). Different ex ante identical individual units of the same type will be scrapped at different final mileage because of variation in mechanical failure and accidents. All of our theory for the representative consumer is robust to allowing for such random product failure—so long as the random failure rates are not endogenous to product taxes—where ϕ_j is interpreted as the mean expected externality from product type j. Once we have an estimate of the average lifetime mileage of each type of vehicle, we multiply it by the vehicle's official government fuel economy rating in order to create a measure of the lifetime carbon externality attached to that product.[8]

We regress these constructed measures of lifetime externalities on vehicle fuel consumption rates and obtain the R^2, which we interpret as the percentage of the first-best welfare gain achieved by the second-best policy (over a baseline of a single tax on all automobiles), and the sum of squared residuals, which we scale by estimates of the own-price derivative of automobiles in order to estimate deadweight loss in dollars. These welfare interpretations rely on the assumptions used in our theory, which we briefly revisit from the perspective of this application before proceeding. First, our model considers a representative consumer—as such, our exercise here abstracts from differences in individual driving behavior.[9] Next, our results rely on assumption 3, which in this context would imply that, conditional on fuel economy, differences in vehicle longevity are uncorrelated with vehicle substitutability. This assumption may not hold exactly, as products of the same brand and similar quality are likely to have more similar durability, so we explore the implications of relaxing that assumption empirically. We find that the initial R^2 appears to be a good approximation of welfare losses even when that assumption is relaxed.

Finally, our results depend upon assumption 1, which states that the externality attached to each product is fixed. In actuality, the externality attached to a particular durable good is endogenous because the decision about when to scrap the product is endogenous.[10] This intensity of use margin is a major concern when considering the debate between a gasoline tax (which provides corrective incentives on the intensity of use margin) versus fuel economy regulations. Our welfare assessments explicitly compare two types of fuel economy tax regimes, and thus this concern is

[8] In doing so, we ignore the carbon emissions related to construction and scrappage of vehicles. Standard estimates suggest that these emissions are only about 10% of total emissions and thus are unlikely to change our results significantly. This also abstracts from 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 not only for simplicity, but also because many climate models and the current federal guidelines suggest that the time path of the social cost of carbon rise at roughly the rate of interest which means that social cost growth offsets discounting.

[9] Note that heterogeneity in annual miles driven is not directly relevant to our analysis here because we care about lifetime mileage of a vehicle. Individual heterogeneity matters to the degree that different individuals depreciate their vehicle faster or slower per mile driven (e.g., because of differences in accident risk or maintenance) or to the degree that their on-road fuel economy for identical vehicles differs due to driving style. See Langer and McRae (2014) for evidence that the variation across consumers in on-road fuel economy is significant, but 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 other models.

[10] See Jacobsen and van Benthem (2015) for evidence on how scrappage decisions influence the welfare implications of fuel economy regulations. Note that this is a particular manifestation of the larger issue related to intensity of use margin. Our concern is over lifetime miles driven of a product, so intensity of use responses (including what is often called the rebound effect (Small and Van Dender 2007)) influence our analysis only to the extent that they imply a change in lifetime utilization, that is to say, only to the degree that they influence scrappage.

largely orthogonal to our main results, but where we interpret our results as commenting on the efficiency of fuel economy regulations (which we do next), this factor implies that our estimates will overstate the efficiency of fuel economy regulations relative to a fuel tax.

Fuel economy policies typically do not directly tax fuel consumption rates, but rather the largest countries in the world all use revenue-neutral fuel-economy standards that impose a shadow price on vehicles that is linear in their fuel consumption rates. In the U.S., the relevant program is the Corporate Average Fuel Economy (CAFE) standard.[11] Historically, that shadow price varied across firms, but the current policy allows cross-firm trading, so that one shadow price should prevail across the market. As such, CAFE creates incentives that are quite similar to the linear tax we model. Our results thus suggest how much more efficient policy could be if the (implicit) tax could depend on not just fuel consumption rates, but also product longevity.

Actually imposing a regulation of this structure, however, would require novel measurement of product durability, and, to the extent that measures of durability are endogenous, it would also induce distortions that attend attribute-based regulation (Ito and Sallee 2015). A gasoline tax would naturally achieve the relative pricing of desired by a tax on fuel consumption and lifetime mileage, to the extent that consumers are aware of product durability and have a rational forward-looking valuation of fuel costs. Thus, our R^2 results can be interpreted as estimates of the fraction of the welfare gain from a gasoline tax that can be achieved by a second-best fuel economy regulation. But, here the caveats mentioned above imply that this comparison overstates the relative efficiency of fuel economy regulation because it ignores the fact that a gasoline tax would also achieve gains along the intensity of use (scrappage) margin. Moreover, a gasoline tax would raise the price of all automobiles, whereas CAFE (implicitly) taxes some products and subsidizes others, and thereby fails to correct the overall size of the automobile market. Taken together, all of this implies that our comparison—within which CAFE is made to look quite poor performing compared to a second-best fuel-economy policy—understates the real welfare losses incurred from using CAFE instead of a tax on gasoline.

3.1 Data

To calculate lifetime mileage for each type of automobile, we use data on vehicle miles traveled (VMT) from California's vehicle emissions testing program—the Smog Check Program—which is administered by the California Bureau of Automotive Repair (BAR). We match the data to a comprehensive registration micro dataset 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. An automobile appears in the data for a number of reasons. First, 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. Second, vehicles more than four years old must pass a smog

[11] Since 2012, the CAFE program adopted standards that are a sliding function of vehicle size. As discussed in Ito and Sallee (2015), this creates additional complications, but such a program still creates a linear shadow price on fuel economy.

check within 90 days of any change in ownership. 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.[12]

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, vintage, and engine characteristics. Using this information, we match the vehicles to Environmental Protection Agency (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. In our main specification, we define each unique 10-digit VIN-prefix ("VIN10-prefix") as a unique vehicle type. This is the finest possible differentiation of ex ante identical vehicles in our data, and it delineates a vehicle according to make, model, model year, engine size and, sometimes, also according to transmission, drive type and body style.

Our primary use of the Smog Check data is to calculate the vehicle's odometer reading shortly before the vehicle was scrapped.[13] However, vehicles may leave the smog check data because they leave California. To accurately determine when a vehicle is scrapped, we also use data obtained from CARFAX Inc. which 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. 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.

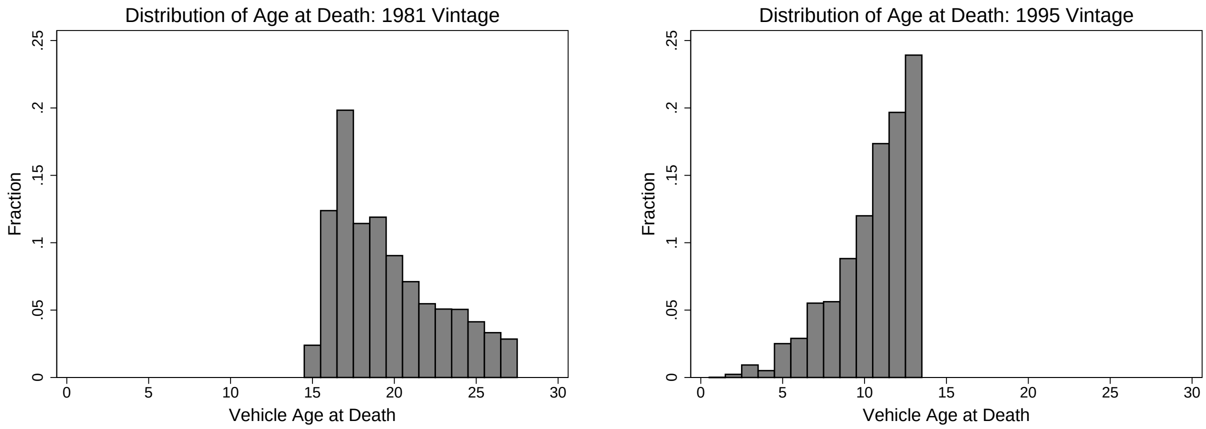
Together, the smog check and CARFAX data give us a measure of the total lifetime vehicle miles traveled for a particular unit, but we do not observe all units, which creates the possibility of measurement error and censorship bias. Regarding the former, we are concerned with the average mileage at scrappage of cars, but we observe only a sample, which could create classical mismeasurement. Classical mismeasurement biases the R^2 from an OLS regression downward (assuming the independent variable is correlated with the true outcome but not the noise). In Section 3.3 we demonstrate that this bias is very small in economic terms for our sample.

Censorship bias is a greater concern because we observe only a subset of all years of retirement for each vehicle type. Vehicles under six years old generally do not appear in the Smog Check data, so we do not observe the lifetime mileage of cars scrapped at very young ages. And, for vehicles that are not yet retired, or were retired before our data began, we do not observe their mileage at scrap, which creates an age censorship that differs across each vintage. For illustration, Figure 2

[12]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.

[13]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.

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



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.[14] This censoring can create (non-classical) mismeasurement, which will be particularly problematic when comparing across cohorts.

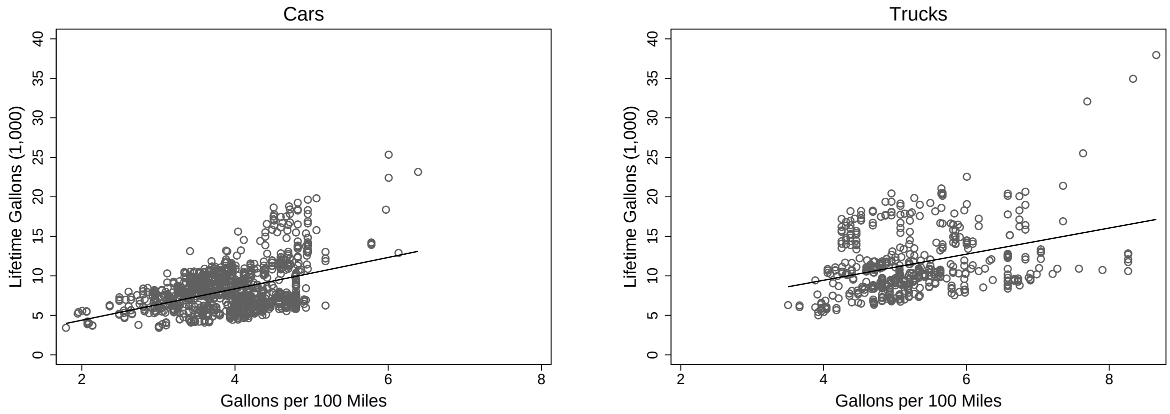
We explore the importance of censoring for our empirical conclusions by comparing how our results vary when we use all of our data versus when we restrict attention to model years 1988 to 1992. As these model years span the age range in which the majority of retirement happens, and censoring of older and younger cars is more or less symmetric, censoring will be less problematic for this subsample. In addition, in Sections 3.2 and 3.3 we show robustness of our results across various samples and extrapolation techniques. We conclude that, while censoring moves our primary estimates, we can bound the impacts of censoring to a sufficiently narrow set of values that have a similar qualitative economic conclusion (which is that ignoring lifetime heterogeneity induces large inefficiencies). For the purpose of extrapolation, we use comprehensive national registration count data from R.L. Polk. These data measure the stock of registered vehicles of each 10-digit VIN-prefix in each year.

3.2 Empirical Results

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 for the relative welfare gain on a tax on fuel economy. Specifically, it is equal to the fraction of the

[14] Our smog check data extend to 2010, but we must observe a two year window after a vehicle's last smog check to know if it has missed its next required check. Thus, we identify vehicle retirements that occurred between 1996 and 2008.

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



Note: The unit of observation is a type of vehicle (a VIN10-prefix). 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.

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 a fuel economy standard.[15] In Section 3.4, 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.

To begin, Figure 3 shows a scatterplot of the relationship between a model's total lifetime externality (gallons of gasoline) and its official fuel consumption rating, for both cars and trucks. A point in the figure corresponds to the average gasoline consumption at the VIN10-prefix level. It 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 substantially 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 from Section 2 hold, this implies that the second-best linear policy captures only 18% and 12% of the

[15]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.

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:

$$\text{Average Lifetime Gasoline Consumption}_j = \alpha + \beta \text{Gallons per Mile}_j + \varepsilon_j, \tag{13}$$

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 an individual retired vehicle.

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.

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 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 above, 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, but not on average product durability, will raise welfare by significantly less than would an efficient policy (such as a carbon tax or a gasoline tax).

Our approach also applies to more flexible fuel economy policies. As long as the policy remains linear in parameters, the R^2 would still be the relevant summary statistic. For example, if the fuel economy policy 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.

Our theoretical results are focused on second-best policies—tax schedules that are set optimally against some design constraint—but real world policies may deviate from the second-best. (By definition, these deviations will lower welfare, so in that sense our estimates of efficiency cost are upper bounds on the actual efficacy of real world policies.) Policies might deviate from the second-best by being "biased" in two different senses. First, policy might get the average tax wrong ("mean bias").[16] Second, policy might have a "slope bias"—the slope of the policy differs from the second-best OLS estimate. One reason that slope bias might emerge is if there is a correlation between average lifetime mileage and fuel consumption ratings in the data, but the policy is determined as if there were no such correlation. In Appendix B we show the difference in this "naïve" tax rate that ignores the correlation between fuel consumption ratings and average lifetime heterogeneity from the actual second-best. In our case, this difference turns out to be small, so we do not emphasize its implications here, though it could be important in other contexts.

In our fuel-economy application, 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). This applies to revenue-neutral fuel-economy standards, which set an average tax rate of zero across all cars.[17] This creates an additional inefficiency at the extensive margin, such that R^2 can be interpreted as an upper bound on the relative welfare gain.

3.3 Robustness: 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 microdata 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 Table B.1 in Appendix B 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, we suspect that most of the cases of reported high lifetime VMT are legitimate. To demonstrate

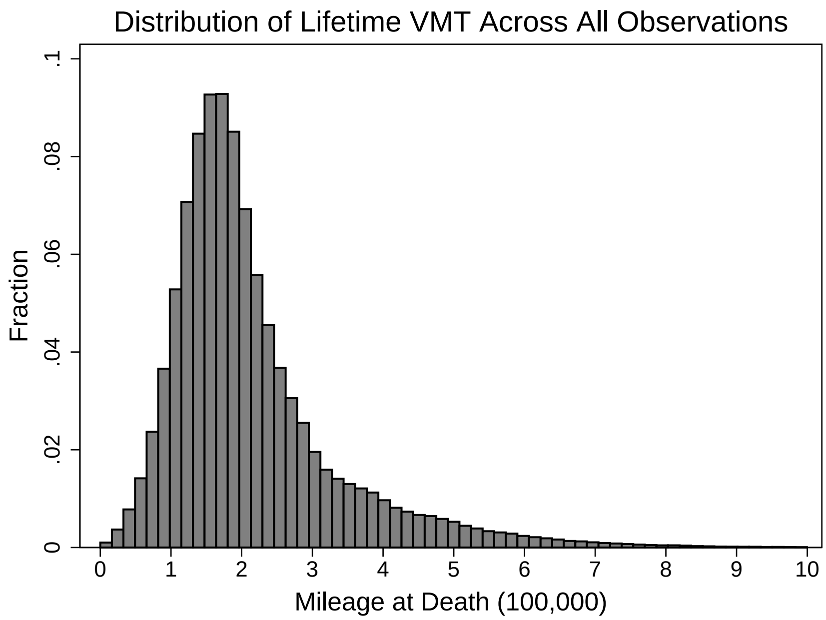
[16]To assess whether fuel economy regulations have mean bias requires an estimate of the shadow price of the regulation, which is beyond the scope of this paper. In terms of the theory, a useful decomposition is to separate the mean bias in tax rates from their variance, which can be seen by rewriting deadweight loss in equation (7) under assumptions 1 to 3:

$$-2 \times DWL(\tau) = \sum_{j=1}^{J} \frac{\partial x_j}{\partial t_j} \times \left(\underbrace{J \times \bar{e}^2}_{\text{bias}} + \underbrace{\sum_{j=1}^{J} (e_j - \bar{e})^2}_{\text{variance}} \right). \tag{14}$$

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. The mean bias can be eliminated by a linear policy, but the variance cannot. Note that OLS minimizes the variance term in equation (14), but (non-OLS) policies with a slope bias have a larger variance term.

[17]Graphically, this can be represented by shifting downward the linear tax schedule in Figure B.1.

Figure 4: Distribution of Lifetime VMT Across All Observations



this, we plot the full histogram of lifetime odometer readings across all of our microdata in Figure 4. 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 facilitates 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 Table 1). Next, we bootstrap the sample again, but in each iteration we randomly drop 50%, 90% or 98% of our sample. 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 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 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 count 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).[18] 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 most missing scrap rates and VMT-at-death are the same across *all* vehicles. This necessarily reduces cross-model variation in lifetime mileage and thus raises the R^2. The process essentially removes all relevant variation for many of 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 presents the results for model years 1988-1992. 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^2s 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 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.

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

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 ≤ 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 3 shows the results for vehicles in the age ranges 10-X years old, where X goes up from 10 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,...,20-year-old cars). The R^2 falls as vehicles get older.[19] Intuitively, 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. 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.

3.4 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.

[19] R^2 decreases from 0.37 for 10-year-old vehicles to 0.13 for 20-year-old vehicles.

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.

Evaluating the formula in equation (5) 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 (on Social Cost of Carbon (2013)), leading to an external cost of 34.6 cents per gallon.[20] Using our data on 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 fuel economy 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 calcula-

[20] If the cost associated with carbon emissions has been rising approximately at the discount rate, we interpret this value as being in 2011 dollars (looking retrospectively at the 1988-1992 vintages).

tion 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 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.

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

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.

to durability induced via the used market as discussed in Jacobsen and van Benthem (2015)) also have the potential to influence welfare.

4 Application 2: Spatial Variation in Emissions

Many externalities have a spatial component, but corrective policies are sometimes unable to differentiate policy treatment over space. For example, regulators worry about the adverse health consequences from high concentrations of ground level ozone (Bell, McDermott, Zeger, Samet, and Dominici 2004). State and federal emissions standards for new vehicles mandate emissions control equipment to reduce local air pollution, including ozone precursors such as NO_x and volatile organic compounds (VOCs). These standards are the same in large parts of the country, yet the marginal damages of these pollutants differ substantially across counties because of differences in population, weather and baseline pollutant concentrations (Muller and Mendelsohn 2009). Another complication is that most miles are driven in areas that have a low concentration of one of the ozone precursors, in which case local emissions in fact do not create ozone pollution (Auffhammer and Kellogg 2011). Hence, the degree to which emissions create an externality depends strongly upon geography, and non-spatially differentiated emissions standards may be poorly targeted.

In this section, we apply our sufficient statistics approach to this type of situation. We suppose that products can be purchased in multiple locations, and the externality that they generate depends on location, but the policy maker can only place a single (e.g., national) tax rate on each good. This is another manifestation of the broad class of problems in which an externality is a function of multiple factors ($\mathcal{A}$) and policy is constrained to be a function of only a subset of those factors ($\mathcal{B}$)—with the key factor in $\mathcal{A}$ that is excluded from $\mathcal{B}$ being geographic location.

When the externality depends on both a product's attributes and its location of use, the errors in a tax policy that depends only on attributes will generally be correlated across products that are used in the same location. Thus, where products used in the same location are closer substitutes (as will generally be the case), assumption 3 (which requires no correlation between relative substitutability and differences in tax errors across products) will likely fail. Thus, in this section, we derive a different sufficient statistic that does not rely on assumption 3, but instead takes advantage of the separability of geographic markets. Our result indicates that the welfare loss of failing to spatially differentiate a policy is summarized by the mean and variance across regions in the externality emitted by identical products. We show that this relates to the R^2 from the second step of a two-stage estimation procedure that removes geographic fixed effects in the first step.

We use this theoretical result to study the efficiency implications of mitigating carbon emissions through the taxation of electricity-consuming home appliances at the national level. In the exercise, we suppose that a product will consume the same amount of electricity regardless of its location, but the carbon emissions generated by electricity consumption vary across location because different fuel sources are marginal in different parts of the electric grid. For example, in California the marginal source of electricity is likely to be natural gas, whereas it will generally be coal in Illinois. We present results specifically for refrigerators, which are large and important home appliances that account for substantial energy usage.[21]

We model a policy in which a planner can establish a tax schedule on refrigerators that differentiates by energy efficiency rating, but the policy must treat a given refrigerator the same regardless of where it is used. While the federal government does not have such a tax policy, we believe that this is a useful characterization of the many existing policies that do indeed treat appliances according to their energy efficiency rating, without taking into account the location of its use. For example, Energy Star certification for appliances and buildings depends solely on energy efficiency, not on location. Many subsidies and rebates for appliances are based on Energy Star certification. Outside of appliances, tax credits for alternative fuel vehicles, weatherization, or solar panels have similar structures. For example, there is a uniform investment tax credit for solar panels, but the greenhouse gas benefits of solar vary with the carbon intensity of the marginal electricity source that it replaces, which varies substantially over space.

4.1 A Modified R^2 as a Sufficient Statistic

To model spatial variation, we suppose that there are $s = 1, ..., S$ geographically distinct markets. We denote the externality (tax) attached to product j consumed in location s by ϕ_{js} (τ_{js}). We model

[21] Refrigerators also offer some modeling simplifications. They have a relatively limited intensity-of-use margin (most refrigerators are run at a constant temperature all of the time), so assumption 1 is plausible. They also have a constant load profile over the diurnal cycle.

the specific case where the externality depends multiplicatively on an attribute of the product f_j and a geography-specific factor r_s: $\phi_{js} = r_s f_j$. The first-best allocation is achievable with Pigouvian taxes, where $t_{js} = r_s f_j$. We suppose that the planner can set a tax on each product that is a linear function of its attribute f_j, but that it cannot not depend on s: $\tau_{js} = \alpha + \beta f_j$.

This captures a wide variety of applications. For example, for SO_2 trading, f_j would be the total SO_2 emissions and r_s would be a factor that scales with population within a radius of the source. Damages depend on both factors, but policy can only place a constant price on industrial activities according to SO_2 emitted. Or, for electricity consumption, f_j is total kilowatt hours consumed and r_s is carbon emissions per kilowatt hour. Damages depend on both factors, but policy can only place a linear tax on products according to electricity consumed.

To proceed, we make assumptions about the nature of the geographic markets. The first is to assume that consumers reside within a particular geographic market and choose among products within that market (i.e., there is no border shopping). This is stated formally in assumption 5:

Assumption 5. *Geographic markets are separable:* $\frac{\partial x_{js}}{\partial t_{kq}} = 0 \ \forall j, k$ *whenever* $q \neq s$.

We believe this to be a minor assumption for large residential purchases because even where bordering shopping does occur, it is feasible to enforce residency-based tax rates. Automobile taxes, for example, are assessed based on state of residence not state of purchase.

Next, for expositional simplicity, we assume that the S markets have identical demand systems, so that cross-derivatives and total demand (conditional on prices) are the same in each market (assumption 6). Finally, we also make use of an assumption that says that the overall size of each product market is fixed (assumption 7).

Assumption 6. *Demand is identical in each geographic market:* $x_{js}(\tau) = x_{jq}(\tau) \ \forall \tau, q, s$.

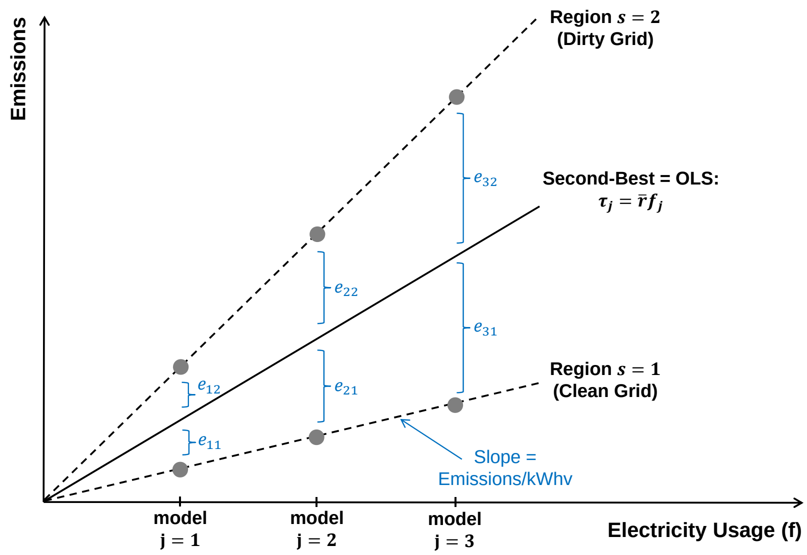
Assumption 7. *Total demand in each geographic market is fixed:* $\sum_j x_{js}(\tau) = \overline{x_s} \ \forall s, \tau$.

Assumption 6 simplifies notation greatly, but is straightforward to relax. Assumption 7 is more substantive, but we can sign the way that relaxing it influences our results. In a discrete choice context—which is often appropriate for energy consuming durable goods—this is akin to assuming there is no net substitution to the outside good. For some products, like refrigerators in the United States, the substitution margin to the outside good is plausibly very small, which makes this assumption appealing. That is, virtually every home has a refrigerator, and few have more than one. Below we demonstrate that welfare results can be bounded when relaxing this assumption.

Under assumption 5, the deadweight loss from imposing a tax vector τ can be written as our original formula in equation (7) summed over the S regions:

$$-2\text{DWL} = \sum_{s=1}^{S} \left(\sum_{j=1}^{J} e_{js}^2 \frac{\partial x_{js}}{\partial t_{js}} + \sum_{j=1}^{J} \sum_{k \neq j} e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \right). \tag{15}$$

Figure 5: Average and Relative Mis-Pricing across Regions



Furthermore, under assumptions 5 to 7, the second-best national tax policy will be to set the policy slope equal to the average damage factor across regions ($\beta = \bar{r}$).[22] The resulting tax error for a product will be the difference between the local damage factor and the average, multiplied by the attribute: $e_{js} = r_s f_j - \tau_{js} = (r_s - \bar{r}) f_j$. Within each region, these second-best taxes create two types of mis-pricing. First, prices are biased (compared to the Pigouvian optimum), on average, across products within each region, depending on how the region's damage factor r_s deviates from the mean. Second, relative prices are also wrong within each region.

This is illustrated in Figure 5, which is a schematic with a hypothetical depiction of three products with different electricity consumption rates in two regions with varying emissions rates. Under the second-best policy, all of the products in region 1 (the clean region) are too expensive, and all of the products in region 2 (the dirty region) are too inexpensive. When there is substitution to an outside good, the bias in each region will create an overall market size distortion—e.g., too many refrigerators are purchased in the dirty market, and too few in the clean market. Within market, the mis-pricing of products is therefore correlated. That is, product 1 in region 2 is under-priced, but so are its substitutes (the other products in region 2). This implies that cross-effects will mitigate own-price effects (the second term of equation (15) partially offsets the first) and the raw R^2 statistic use in Proposition 1 would overstate the inefficiency.

The second type of mis-pricing is that even within-region, relative prices are wrong. Figure 5 demonstrates that the slope of the second-best OLS tax schedule does not equal the slope in any of the two regions. Hence, even if the OLS policy could be adjusted for each region to get the average tax rate correct, products with similar attributes have similar structural tax errors.

Assumption 7 implies that the first type of mis-pricing—average bias within a region—creates

[22] A proof of this is included in the proof of Proposition 3 in Appendix A.

no distortion in choice.[23] In this case, a simple sufficient statistic is available that captures the relative efficiency of a national linear policy as compared to the efficient spatially-differentiated tax, evaluated over a baseline of a flat unbiased tax on all products. This result is stated in Proposition 3 (see Appendix A for a proof).

Proposition 3. *Under assumptions 1, 2, 5, 6 and 7, the second-best policy is $\beta = \bar{r}$, and the fraction of the first-best welfare gain achieved by this second-best policy over a policy of a constant unbiased tax on all products is:*

$$\frac{DWL(\tau = \bar{r} f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} = 1 - \frac{var(r_s)}{E[r_s^2]}. \tag{16}$$

Note that this result depends only upon the variation in damage factors across regions. The result does not rely on any assumptions about the demand system (e.g., the second-type of mis-pricing mentioned above is irrelevant), except that the geographic markets are separable and there is no net substitution to the outside good. The structure of demand within each region will influence the deadweight loss measured in dollars[24], but variation in the demand system affects welfare under the relevant policy alternatives in a proportional way, so that all the demand derivatives divide through.

As with our results in Section 2, this result relates to familiar statistics from regression analysis. In particular, the proportional welfare gain expressed in Proposition 3 is equal to the R^2 from a regression of regionally demeaned damages ϕ_{js} on the product attribute. This is stated in corollary 2, which uses the object $\bar{f}$, which is the sales-weighted average value of the attribute f_j (proof in Appendix A).

Corollary 2. *Under assumptions 1, 2, 5, 6 and 7, the fraction of the first-best welfare gain achieved by the second-best policy is equal to the R^2 from a regression of $\phi_{js} - r_s \bar{f}$ on f_j.*

Intuitively, when there is no net substitution to the outside good, then the regional biases in each market do not influence welfare, but the within-region relative mispricing does create distortions. Demeaning the externality data within each region removes these market size effects, leaving only the relevant variation in externalities.

As noted above, the assumption of no net substitution to the outside good will be plausible in some conditions, but not others. When the j goods represent all of the goods in a sector, it is logical to assume that, on average across products, increases in taxes on each product will lead to a decrease in the total market size (i.e., the sector is not a “Giffen sector”, where average price

[23] The constant α in a tax schedule $\tau_{js} = \alpha + \beta f_j$ is therefore irrelevant. α does affect a transfer between consumers and the government, but this is undone through revenue recycling.

[24] As shown in the proof of Proposition 3 in Appendix A, the deadweight loss expression can be obtained by simplifying the expression in equation (15):

$$-2\text{DWL} = S \times \text{var}(r_s) \times \sum_{j=1}^{J} \sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}. \tag{17}$$

increases expand the market). Under that assumption, the welfare statistic derived above will understate the fractional welfare gain of the second-best policy over the baseline of an equal tax on all products, which we state in corollary 3.[25]

Corollary 3. *Under assumptions 1, 2, 5, and 6, the second-best policy is $\beta = \bar{r}$, and the fraction of the first-best welfare gain achieved by this second-best policy over a policy of a constant tax on all products is:*

$$\frac{DWL(\tau = \bar{r} f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} > 1 - \frac{var(r_s)}{E[r_s^2]}. \tag{18}$$

Thus, the same sufficient statistic represents a lower bound on the fraction of the first-best welfare gain that is achieved by the non-spatially differentiated product tax when there is substitution to the outside good. In our empirical application, which we move to next, the sufficient statistic is quite large, and thus it is an informative lower bound on the fractional welfare gain.

4.2 Data and Results

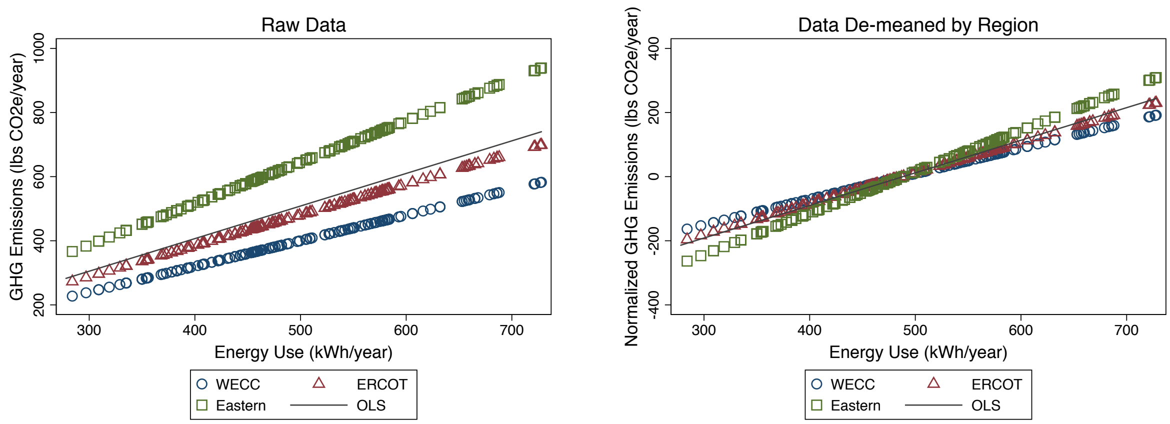
We use Proposition 3 to evaluate the welfare properties of a national tax on the energy efficiency of refrigerators as a carbon mitigation policy. We choose refrigerators because they are a major household appliance, and their electricity consumption is particularly easy to model. For this purpose, we obtained the energy-efficiency rating for a cross-section of all refrigerators certified for sale in the United States in 2010 from the Association of Home Appliance Manufacturers. We exclude compact refrigerators. Our final sample includes 1,349 models. Their average government rated electricity consumption is 488 kWh per year, with a standard deviation of 93 kWh.

Deploying two identical refrigerators in two different locations will have different emissions implications when their marginal electricity demand is met by increased production from power plants with different emissions rates. Power markets are integrated over geographic regions, so differences in emissions will emerge mostly across these markets. The power market is thus the relevant unit of spatial heterogeneity for our analysis. Existing literature suggests that the appropriate level of integration is to either consider the three major power market interconnections, or to consider eight distinct regions defined by the North American Electric Reliability Corporation (NERC) (Graff Zivin, Kotchen, and Mansur 2014; Holland et al. 2014).

To quantify spatial differences in emissions rates, we rely on results from Graff Zivin, Kotchen, and Mansur (2014), who estimate the emissions rate from the marginal generation of electricity at each hour of the day in each of several electricity regions. We assume that refrigerators use a constant level of electricity throughout the day, so we simply take the average over these marginal rates over the 24 hour cycle. Multiplying the average marginal emissions rate by the energy consumption rate per year of each refrigerator yields an estimate of the carbon emissions per year that

[25]The proof is in Appendix A. For expositional clarity, the proof uses an additional regulatory condition, which says that average substitution to the outside good across products is not correlated with the attribute and with demand derivatives.

Figure 6: Distribution of Annual Carbon Emissions for Refrigerators by Power Market



Note: Graph shows a ten percent random sample of all refrigerator models for visual clarity. The three regions modeled are the Eastern Interconnection (Eastern), the Electric Reliability Council of Texas (ERCOT) and the Western Electricity Coordinating Council (WECC).

would be expected for a product deployed in each electricity region. We use this annual emissions rate in our regressions.[26]

The left panel in Figure 6 shows a scatterplot of annual emissions against annual electricity consumption for a sample of products across the three power market interconnections. Within a region, there is, by construction, a perfect linear relationship between fuel consumption and emissions. However, both the level and the slope differ across regions because of differences in emissions rates per kWh. The figure also shows the OLS fit to the raw data. As discussed above, the OLS slope is an average of the slopes across the three regions; this is the second-best policy under the assumptions used in Proposition 3. Residuals are substantial and are highly correlated across regions. Thus, if the OLS line represented a tax schedule, there would be substantial errors, but those errors are similar (though not identical) across products within a region.

The right panel in Figure 6 shows the regionally demeaned data. These are the data that enter our modified R^2 statistic. The residuals are greatly muted. If the OLS line was a tax schedule, there is still a relative mis-pricing of products within a region because the slope of the tax function is too steep for the power markets with lower than average emissions rates, whereas it is not steep enough for the markets with higher than average emissions rates.

Table 5 shows the R^2 values from a regression of the carbon emissions associated with each product by region observation on the product's electricity consumption rate. The top line regresses the region demeaned data. As shown in Corollary 2, these R^2 values, which are 0.96 for the case of three interconnections and 0.90 for the eight NERC regions, represent the fraction of the first-best welfare gain achieved by a national policy that does no spatial differentiation over a baseline policy

[26] All values could be rescaled by a common factor to convert annual emissions in lifetime emissions (assuming a common lifespan for each model); this will have no effect on the R^2 because it amounts to scaling the dependent variable.

Table 5: Sufficient Statistics for Spatially Differentiated Refrigerators

	Three interconnections	**Eight NERC regions**
R^2 from demeaned data	0.96	0.90
R^2 from raw data	0.47	0.24
Sample size	4,047	10,792

Note: Table shows R^2 from a regression of demeaned emissions (first row) or actual emissions (second row) on electricity consumption. The unit of observation is a refrigerator model in a particular interconnection (second column) or NERC region (third column).

that taxes all refrigerators a flat amount. For reference, the R^2 values from the raw data, which are far smaller, are also included.

These findings demonstrate a, perhaps surprisingly, small welfare loss of the lack of regional policy differentation for electric appliances. The estimates from the raw data show that spatial differences in emissions rates per kWh do create large differences in implied emissions for products sold across parts of the United States. But, the differences are largely *between* geographic markets. The resulting mis-pricing of products within a market from a national policy are highly correlated within a geographic market, which mutes the degree to which mis-pricing causes consumers to choose the wrong appliance. As a result, the welfare impacts of failing to spatially differentiate corrective taxes across electricity markets are modest (relative to the total gain achieved by optimal policy over the baseline). This result relies on the extensive margin for refrigerator demand being zero or small. Where that extensive margin response grows (and it will be larger for other appliances than refrigerators), there will be a second welfare loss due to the fact that the overall product market will be too large or small in each region. And, as shown in Corollary 3, this implies that the already high R^2 values understate the fraction of the welfare gain achieved by the national policy.

5 Application 3: Noisy Energy Efficiency Ratings

Another reason that taxes or regulatory incentives for energy-consuming products may be imperfectly related to the true externalities that they generate is that the energy efficiency ratings themselves are imperfect. To determine the energy efficiency rating of a product, governments establish a laboratory test procedure. The government, or the manufacturers themselves, then test a prototype or example product. Actual performance in the field can differ from lab test results and, when it does, policies based upon the official ratings will be imperfect indicators of the actual externalities associated with each product. This creates inefficiencies that are similar to the other applications described above, and our sufficient statistics framework can be used to quantify the consequent welfare losses.

In general, the challenge in studying this phenomenon is that it requires credible measures of average in-use energy efficiency which can be compared to the official rating. Scattered evidence of in-use performance do exist for some products, but we take a different approach here and analyze a change in the U.S. rating system for automobiles that was meant to address mis-measurement. The EPA began measuring fuel economy of automobiles in 1978 in support of the CAFE program. The ratings are based on a laboratory test during which a vehicle is driven on a dynamometer (a treadmill for cars) through a specific pattern of speeds and accelerations.[27] The test procedure established in 1978 included two courses; one was to represent urban driving, and the other represented highway travel. The two ratings were averaged to determine each vehicle's rating for the CAFE program. These same ratings were presented to consumers on fuel economy labels.

In 1986, in response to consumer complaints that the ratings systematically overstated fuel economy, the EPA revised the ratings downward by simply scaling them by the same amount for all vehicles. City values were reduced by 10%, and highway values were reduced by 22%. CAFE continued to use the original values to determine automakers' compliance, but consumer labels were updated. Over time, the revised ratings were deemed to be inaccurate as well. The original test used low highway speeds, did not involve the use of air conditioning, and generally became less accurate as automobile technology and average driving patterns changed. Yet again, the EPA instituted a new test procedure that changed the ratings substantially on average, and also more for some vehicles than others.[28]

For political reasons, however, the CAFE program continues to use the less accurate original rating system from 1978.[29] While consumers are now provided with the more accurate updated ratings, the regulation (and hence the regulatory shadow price faced by automakers) are still based on the noisy original system.

We can use our sufficient statistics framework to quantify the welfare costs of using the old rating system. Our thought experiment is the following. We suppose (1) the new ratings represent the true fuel economy rating of a vehicle, (2) after a linear adjustment, the old rating is a white noise mismeasurement of the truth, and (3) the policy maker is sophisticated and is aware of the inaccuracy in the old rating but must base policy upon it because of political or legal constraints. These assumptions likely hold in practice—reported on-road fuel economy is close to the 2008 EPA ratings on average and the EPA explicitly presents differences between window sticker and regulatory CAFE fuel-economy values.[30] [31]

For a standard fuel economy policy that imposes a linear shadow price, the tax schedule will be linear in the old fuel consumption rating, labeled F_{old}: $\tau_j = \alpha + \beta F_{old}$.[32] The errors in the tax

[27]Emissions from the vehicle are captured and measured in order to determine the fuel consumed. This procedure is more reliable than measuring the volume of the liquid fuel before and after the test. The chemical composition of the fuel input is known, so emissions provide a precise measurement of gasoline consumption.

[28]This procedure involved five separate dynamometer tests – the original two tests and three new ones. Several tests are combined together to determine the highway and city ratings that appear on fuel economy labels for consumers.

[29]Evidently it was determined that changing the rating that entered the CAFE compliance program would require a political battle not worth waging. Changing the CAFE ratings would have created winners and losers among automakers.

[30]See http://www.epa.gov/fueleconomy/documents/420f14015.pdf and http://www.epa.gov/fueleconomy/documents/420b14015.pdf.

[31]If the 2008 ratings are still noisy, R^2 is likely an upper bound on the efficiency gain from the second-best policy based on 1978 ratings, unless the remaining noise is structurally correlated with substitutability across vehicle models.

[32]We specify this in terms of fuel consumption, the inverse of fuel economy, so that total externality scales linearly

rate are then the difference between this tax schedule and the actual externality, which is some scalar (which represents the externality per gallon of gasoline consumed times the miles driven) multiplied by the new rating: ψF_{new}.[33] The error is then $e_j = \psi F_{new} - \alpha - \beta F_{old}$.

Where the noise in measurement is uncorrelated with factors that determine vehicle demand, assumption 3 will hold, and the main theoretical results in Proposition 1 and Corollary 1 apply. It is logical to suppose that cross-effects would cancel, as errors in the tax rates from an unbiased policy are likely to be due to particular technologies, like stop-start systems, that are of little concern to consumers (and therefore not correlated with cross-price derivatives) or idiosyncratic aberrations from test trials. In this case, the R^2 from a regression of F_{new} on F_{old} represents a sufficient statistic. As usual, it indicates the fraction of the welfare gain over a flat tax (that corrects for the average externality produced by an automobile) achieved by a policy that uses the less accurate, noisy fuel economy estimates in place of the accurate ratings.

5.1 Data and Results

To estimate this R^2, we use the sample of vehicles that the EPA itself used to establish the concordance between the old and new highway and city test ratings. In determining how to create the new system, the EPA tested a few hundred vehicles meant to represent the car market and compared the results under the new and old regimes. We obtained the data from these tests from the EPA and use them here to assess the change in ratings.[34]

We plot these data in Figure 7. The old and new ratings are highly correlated, but there is an upward bias in the old ratings (the old miles per gallon ratings were too high on average). In addition, there are noticeable differences in how the test revision affected different models—there is dispersion around the fitted line. The rating change is quantitatively important: the average difference between the old and new estimated present-discounted fuel costs in this sample is $1,700. The difference ranges from $500 to $4,250 with a standard deviation of nearly $700.[35] So even if the bias was recognized, it still affected different vehicles to varying degrees.

The OLS regression of the new rating on the old one yields an R^2 above 0.97.[36] This indicates that the vast majority of the welfare gain from an optimally designed fuel-economy policy that used the new ratings is achieved by a policy that uses the old rating system. Interestingly, this makes the lack of updating relatively innocuous despite the fairly large differences between the two rating systems. The welfare losses from this noise, however, may be substantial if the policy maker does not take the bias in the old ratings into account and fails to make a correction (i.e., chooses

with the rating.

[33] In this thought experiment, we abstract away from the heterogeneity in longevity that we emphasize in Section 3 above, and assume that ψ is the same (in expectation) for all vehicles.

[34] These same data are used in Sallee (2014) to characterize the uncertainty faced by consumers about true lifetime fuel costs of vehicles under the old regime.

[35] This assumes a gasoline price of $2.50 per gallon (roughly the average in 2008), for vehicles driven 12,000 miles per year for 14 years with a 5% discount rate.

[36] The R^2 changes little when modifying the sample. Adding the 13 available hybrid models to the gasoline-powered sample produces an R^2 of 0.98. The R^2 values for the subsamples of cars and trucks are 0.96 and 0.98, respectively.

Figure 7: Old and New Combined Fuel Economy Ratings

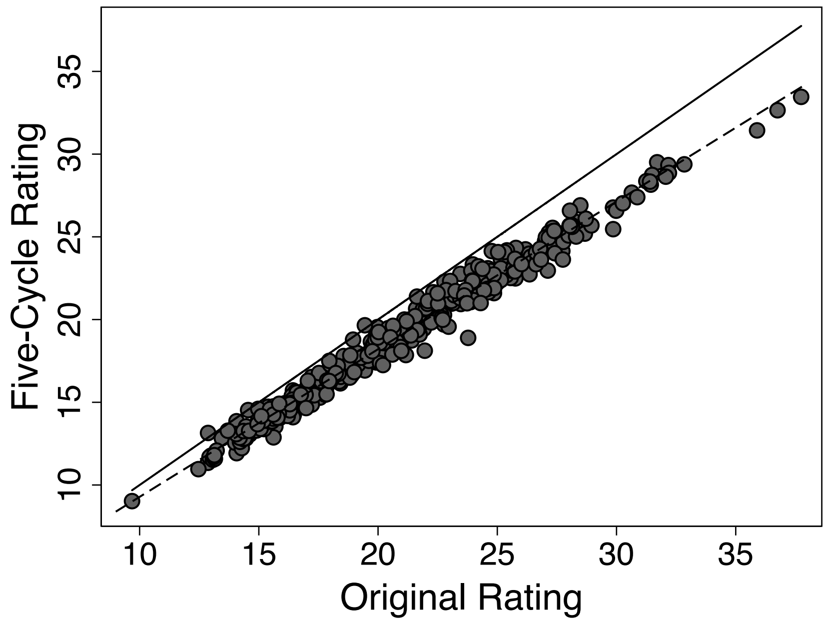


Figure shows the pre-2008 (original) combined fuel economy rating and the post-2008 (five-cycle) rating in miles per gallon for a sample of vehicles. Dashed line is linear fit. Solid line is the 45-degree ray.

a policy that is based on the assumption that the old rating system is accurate and is therefore too lax on average, causing distortions on the extensive margin). Also, note that the inefficiency from noisy energy efficiency ratings adds to a long list of existing distortions from fuel-economy standards, including the welfare loss from ignoring product longevity discussed in Section 3.[37]

6 Conclusion

Externality-correcting policies rarely take on the ideal form of a direct tax on marginal damages. Actual policies are frequently constrained by administrative feasibility or political constraints so that they must place imperfect marginal incentives on products or actions. This paper provides a framework for identifying sufficient statistics that facilitate evaluation of the welfare costs of such constraints that require limited market information. We emphasize that, under certain conditions, simple regression statistics have welfare interpretations.

The paper demonstrates the usefulness of this approach through three examples. The examples are all related to environmental externalities, but they span a number of distinct challenges to policy, including random mismeasurement of damages, spatial heterogeneity, and the implications of heterogeneity in the lifetime utilization of energy-consuming durable goods. These applications

[37]Note that real-world driving patterns may deviate from the new test procedures as well. Data on in-use performance of individual models could be used to test for this. Similar exercises could be done for other durables.

demonstrate the viability of our theoretical framework, but they also make contributions in their own right.

Most importantly, our study of the heterogeneity in automobile longevity points out a previously undiscussed efficiency problem with a class of energy efficiency policies that regulate new durable goods. When different products have different average lifetime utilization, energy efficiency policy—which creates explicit or implicit price incentives according to only energy efficiency ratings—is inherently imprecise. Through analysis of unique micro data on automobile mileage, we demonstrate that different types of automobiles have widely varying average lifetime mileage, which implies large inefficiencies in fuel economy policy, on the order of $2 billion per year, that result from the relative mis-pricing of automobiles.

References

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.

Auffhammer, Maximilian and Ryan Kellogg. 2011. "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality." *American Economic Review* 101 (6):2687–2722.

Bell, Michelle L., Aidan McDermott, Scott L. Zeger, Jonathan M. Samet, and Francesca Dominici. 2004. "Ozone and Short-Term Mortality in 95 US Urban Communities, 1987-2000." *Journal of the American Medical Association* 292 (19):2372–2378.

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

Borenstein, Severin. 2015. "A Microeconomic Framework for Evaluating Energy Efficiency Rebound and Some Implications." *Energy Journal* 36:1–21.

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.

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.

Fowlie, Meredith and Nicholas Z. Muller. 2013. "Market-based Emission Regulation when Damages Vary Across Sources: What are the Gains from Differentiation?" Manuscript: UC Berkeley.

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.

Graff Zivin, Joshua S., Matthew Kotchen, and Erin T. Mansur. 2014. "Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies." *Journal of Economic Behavior and Organization* 107 (A):248–268.

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

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.

Ito, Koichiro and James M. Sallee. 2015. "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards." Manuscript: University of Chicago.

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." URL http://www.nber.org/papers/w18849. NBER Working Paper No. 18849.

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

Mendelsohn, Robert. 1986. "Regulating Heterogeneous Emissions." *Journal of Environmental Economics and Management* 13:301–312.

Muller, Nicholas Z. and Robert Mendelsohn. 2009. "Efficient Pollution Regulation: Getting the Prices Right." *American Economic Review* 99 (5):1714–1739.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus. 2011. "Environmental Accounting for Pollution in teh United States Economy." *American Economic Review* August (5):1649–1675.

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

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.

———. 2014. “Rational Inattention and Energy Efficiency.” *Journal of Law and Economics* 57 (3):781–820.

Small, Kenneth A. and Kurt Van Dender. 2007. “Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect.” *Energy Journal* 28 (1):25–51.

A Appendix: Proofs

Derivation of Equation (5)

Let any generic tax schedule be denoted as $\tau_1, ..., \tau_J$. To obtain equation (5), we characterize the welfare loss of moving from the optimal 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 ϕ_j to τ_j. That is, we specify the function $t_j = (1-\rho)\phi_j + \rho\tau_j$. We differentiate W with respect to ρ, and then derive the welfare loss of moving from the optimal policy (when $\rho = 0$) to the alternative policy (when $\rho = 1$).

First, we differentiate equation (2) with respect to ρ and substitute in the consumer's optimality condition. This yields:

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

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. 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).

By definition, $\frac{\partial t_k}{\partial \rho} = (\tau_k - \phi_k)$. We use that substitution, as well as the definition of t_j, and simplify:

$$\begin{aligned}\frac{dW}{d\rho} &= \sum_{j=1}^{J}\sum_{k=1}^{J}\left(\{(1-\rho)\phi_j + \rho\tau_j\} - \phi_j\right)\frac{\partial x_j}{\partial t_k}(\tau_k - \phi_k) \\ &= \rho\sum_{j=1}^{J}\sum_{k=1}^{J}(\tau_j - \phi_j)\frac{\partial x_j}{\partial t_k}(\tau_k - \phi_k).\end{aligned} \tag{A.2}$$

Because ρ is a constant, we can remove it from the summation, which yields the final equation. To obtain the change in social surplus from moving fully between the two tax schedules, we integrate from $\rho = 0$ to $\rho = 1$. If the demand derivatives (assumption 2) are constant over the relevant range, then the integration is straightforward and yields:

$$W(t = \phi) - W(t = \tau) \equiv DWL(\tau) = -\frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J}(\tau_j - \phi_j)(\tau_k - \phi_k)\frac{\partial x_j}{\partial t_k}. \tag{A.3}$$

Proof of Proposition 1

First, differentiate equation (8) (the expression for deadweight loss) with respect to α and β to obtain first-order conditions. For α, this yields:

$$\frac{\partial DWL}{\partial \alpha} = -\frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} -(\phi_j - \alpha - \beta f_j)\frac{\partial x_j}{\partial t_k} - (\phi_k - \alpha - \beta f_k)\frac{\partial x_j}{\partial t_k} = 0. \tag{A.4}$$

Quasilinearity implies symmetry of the demand derivatives, so separating the two terms and summing over all j and k yields the same quantity. Combining the terms and dividing through by constants yields:

$$\sum_{j=1}^{J}\sum_{k=1}^{J}(\phi_j - \alpha - \beta f_j)\frac{\partial x_j}{\partial t_k} = 0. \tag{A.5}$$

Solving for α yields:

$$\begin{aligned}\alpha &= \left(\sum_{j=1}^{J}\sum_{k=1}^{J}\frac{\partial x_j}{\partial t_k}\right)^{-1}\sum_{j=1}^{J}\sum_{k=1}^{J}(\phi_j - \beta f_j)\frac{\partial x_j}{\partial t_k} \\ &= \sum_{j=1}^{J}\sum_{k=1}^{J}(\phi_j - \beta f_j)\omega_{jk},\end{aligned} \tag{A.6}$$

where $\omega_{jk} \equiv (\frac{\partial x_j}{\partial t_k})/(\sum_{j=1}^{J}\sum_{k=1}^{J}\frac{\partial x_j}{\partial t_k})$ is a weighting parameter.

The same steps (including use of symmetry) for β yields:

$$\sum_{j=1}^{J}\sum_{k=1}^{J} f_j(\phi_k - \alpha - \beta f_k)\frac{\partial x_j}{\partial t_k} = 0. \tag{A.7}$$

Substituting in for α from equation A.6 and solving for β yields a closed-form expression for the second-best β:

$$\beta = \frac{\sum_{j=1}^{J}\sum_{k=1}^{J} f_j\phi_k\omega_{jk} - \sum_{j=1}^{J}\sum_{k=1}^{J}\phi_k\omega_{jk}\sum_{j=1}^{J}\sum_{k=1}^{J} f_j\omega_{jk}}{\sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k\omega_{jk} - \sum_{j=1}^{J}\sum_{k=1}^{J} f_k\omega_{jk}\sum_{j=1}^{J}\sum_{k=1}^{J} f_j\omega_{jk}}. \tag{A.8}$$

We then substitute in the definition of the structural error, $\varepsilon_j \equiv \phi_j - \alpha^{OLS} - \beta^{OLS} f_j$ to eliminate all ϕ terms. Rearranging this yields:

$$\begin{aligned}\beta =&\beta^{OLS} \frac{\sum_{j=1}^J \sum_{k=1}^J f_j f_k \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J f_k \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}}{\sum_{j=1}^J \sum_{k=1}^J f_j f_k \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J f_k \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}} \\ &+ \alpha^{OLS} \frac{\sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}}{\sum_{j=1}^J \sum_{k=1}^J f_j f_k \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J f_k \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}} \\ &+ \frac{\sum_{j=1}^J \sum_{k=1}^J f_j \varepsilon_k \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J \varepsilon_k \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}}{\sum_{j=1}^J \sum_{k=1}^J f_j f_k \omega_{jk} - \sum_{j=1}^J \sum_{k=1}^J f_k \omega_{jk} \sum_{j=1}^J \sum_{k=1}^J f_j \omega_{jk}} \\ =&\beta^{OLS}. \end{aligned} \tag{A.9}$$

The last step follows from several simplifications. First, the term multiplying β^{OLS} is equal to 1 (numerator and demonimator are the same). Note that $\sum_{j=1}^J \sum_{k=1}^J \omega_{jk} = 1$ by definition of ω_{jk}, which is a weight. Thus, the numerator in the term multiplying α^{OLS} is zero.

Finally, by assumption 3, the numerator of the final term in line (A.9) is also zero. Multiplying out to rewrite the ω terms as derivatives, the first term in the numerator can be written as $\sum_{j=1}^J f_j \sum_{k=1}^J \varepsilon \frac{\partial x_j}{\partial t_k}$. This is decomposed into own and cross effects as: $\sum_{j=1}^J f_j \varepsilon_j \frac{\partial x_j}{\partial t_j} + \sum_{j=1}^J f_j \sum_{j \neq k} \varepsilon_k \frac{\partial x_j}{\partial t_j}$. The cross effects are zero directly from assumption 3. The own-terms are zero because ε_j and f_j are uncorrelated (by construction, from OLS); ε_j and $\frac{\partial x_j}{\partial t_j}$ are uncorrelated (assumption 3); and f_j and $\frac{\partial x_j}{\partial t_j}$ are uncorrelated (assumption 4). The own terms thus equal $J \times \bar{f} \bar{\varepsilon} \overline{\frac{\partial x_j}{\partial t_j}} = 0$ because $\bar{\varepsilon} = 0$. The second term in line (A.9) is also zero. Rewrite it as: $\sum_{j=1}^J \sum_{k=1}^J \varepsilon_j \frac{\partial x_j}{\partial t_k} = \sum_j \varepsilon_j \frac{\partial x_j}{\partial t_j} + \sum_{j=1}^J \sum_{k \neq j} \varepsilon_j \frac{\partial x_j}{\partial t_k}$. Both these terms are zero, directly from assumption 3.

This is zero for each j by assumption 3, which says there is no correlation between the structural errors and the demand matrix. Likewise, rewriting the first summation in the second term yields $\sum_{j=1}^J \sum_{k=1}^J \varepsilon \frac{\partial x_j}{\partial t_k}$, which is again zero by the assumption of no correlation between errors and the demand matrix.

Having obtained the second-best β, we can substitute this back into equation A.6, along with the structural error definition:

$$\begin{aligned}\alpha &= \sum_{j=1}^J \sum_{k=1}^J (\alpha^{OLS} + \beta^{OLS} f_j + \varepsilon_j - \beta^{OLS} f_j) \omega_{jk} \\ &= \alpha^{OLS} \sum_{j=1}^J \sum_{k=1}^J \omega_{jk} + \sum_{j=1}^J \sum_{k=1}^J \varepsilon_j \omega_{jk} = \alpha^{OLS}. \end{aligned} \tag{A.10}$$

Again, summing over the weights yields 1, and assumption 3 implies that the second term is zero. This shows that the second-best α and β are equal to the OLS parameters α^{OLS} and β^{OLS}.

Given that, we can then evaluate the deadweight loss formula by substituting in these second-best values into equation (7), which decomposes own and cross effects:

$$\begin{aligned} DWL &= -\frac{1}{2}\sum_{j=1}^{J}\varepsilon_j^2\frac{\partial x_j}{\partial t_j} - \frac{1}{2}\sum_{j=1}^{J}\sum_{k\neq j}\varepsilon_j\varepsilon_k\frac{\partial x_j}{\partial t_k} \\ &- \frac{1}{2}\sum_{j=1}^{J}\varepsilon^2\frac{\partial x_j}{\partial t_j} \\ &= -\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}SSR \end{aligned} \tag{A.11}$$

Under assumption 3 the cross-effects terms are zero, which produces the first simplification. The third equality follows from assumption 4, which says that own-derivatives are uncorrelated with . As the ε terms are the OLS residuals, the summation is, by definition, the sum of squared residuals from OLS.

Proof of Corollary 1

This follows from Proposition 1, which states that DWL $= -\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}SSR$, and the expression for the total sum of squares from the OLS regression of ϕ on f:

$$TSS = \sum_{j=1}^{J}(\phi_j - \bar{\phi})^2 = \sum_{j=1}^{J}(\phi_j - \alpha^{OLS})^2, \tag{A.12}$$

where α^{OLS} is the estimate from an OLS regression on a constant only. $\alpha^{OLS} = k$ equals the second-best constrained constant tax: the average externality. Therefore, under assumptions 1 to 4, $DWL(\tau = k) = -\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}TSS$. Now we can evaluate the equation (11) in Corollary 1 as:

$$\begin{aligned} \frac{DWL(\tau = \alpha^{OLS} + \beta^{OLS}f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} &= \frac{-\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}SSR - (-\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}TSS)}{0 - (-\frac{1}{2}\overline{\frac{\partial x_j}{\partial t_j}}TSS)} \\ &= \frac{TSS - SSR}{TSS} = \frac{ESS}{TSS} = R^2. \end{aligned} \tag{A.13}$$

Proof of Proposition 2

This follows directly from Proposition 1 and the fact that equation (8) can be written in matrix notation as:

$$\min_{\mathbf{b}} DWL(\tau) = -\frac{1}{2}\mathbf{e'De} \tag{A.14}$$

where $\mathbf{D}$ is the matrix of own- and cross-price derivatives of demand, the vector $\mathbf{e} = \boldsymbol{\phi} - \mathbf{Fb}$ in which $\boldsymbol{\phi}$ is the vector of product-specific externalities, $\mathbf{F}$ is the matrix of product attribute values including a constant, and $\mathbf{b} = (\alpha\ \beta)'$ is the vector of policy coefficients. Now redefine $\mathbf{D^*} = -\mathbf{D}$, so that the problem becomes to minimize $\mathbf{e'D^*e}$. This is exactly the definition of a generalized least squares estimation.[38]

Proof of Proposition 3

We first show that under assumptions 1, 2, 5, 6 and 7, the second-best policy is $\alpha = 0$ and $\beta = \bar{r}$. Consider an OLS policy $\tau_{js} = \alpha + \beta f_j$. Residuals are given by $\phi_{js} - (\alpha + \beta f_j) = (r_s - \beta) f_j - \alpha$, which uses the fact that $\phi_{js} = r_s f_j$. By equation (15), the deadweight loss from this OLS policy is:

$$
\begin{aligned}
-2\text{DWL}(\tau = \alpha + \beta f_j) &= \sum_{s=1}^{S}\sum_{j=1}^{J}\sum_{k=1}^{J} e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^{S}\sum_{j=1}^{J}\sum_{k=1}^{J} ((r_s - \beta) f_j - \alpha)((r_s - \beta) f_k - \alpha) \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^{S} (r_s - \beta)^2 \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \alpha \sum_{s=1}^{S} (r_s - \beta) \sum_{j=1}^{J}\sum_{k=1}^{J} (f_j + f_k) \frac{\partial x_{js}}{\partial t_{ks}} \\
&+ \alpha^2 \sum_{j=1}^{J}\sum_{k=1}^{J} \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \times \sum_{s=1}^{S} (r_s - \beta)^2, \qquad \text{(A.15)}
\end{aligned}
$$

where the third equality follows from assumption 6 (common demand system in each market), which implies that $\frac{\partial x_{js}}{\partial t_{ks}} = \frac{\partial x_{jq}}{\partial t_{kq}}\ \forall q, s$. The fourth equality follows from the two facts. First, under assumption 7 (no substitution to the outside good), $\sum_{j=1}^{J}\sum_{k=1}^{J} \frac{\partial x_{js}}{\partial t_{ks}} = 0$ so the final term (with α^2) is zero. Second, under quasilinearity, the demand matrix is symmetric, so $\sum_{j=1}^{J}\sum_{k=1}^{J} f_j \frac{\partial x_{js}}{\partial t_{ks}} = \sum_{j=1}^{J}\sum_{k=1}^{J} f_k \frac{\partial x_{js}}{\partial t_{ks}} = 0$, and under assumption 7 (no substitution to the outside good), these terms are equal to zero: $\sum_{j=1}^{J}\sum_{k=1}^{J} f_j \frac{\partial x_{js}}{\partial t_{ks}} = \sum_{j=1}^{J} f_k \sum_{k=1}^{J} \frac{\partial x_{js}}{\partial t_{ks}} = \sum_{j=1}^{J} f_k \times 0 = 0$. Thus, both terms involving α are equal to zero, and deadweight loss is reduced to only the first term involving β.

The α terms cancel because α is just a lump sum transfer between the government and consumers. Given assumption 7, with revenue-recycling, the constant has no effect on welfare and the optimal α is undetermined. We set it to zero, which makes the tax rate an unbiased estimate of the externality. To find the second-best policy, we minimize deadweight loss (maximize expression A.15) with respect to β. The first-order condition is:

[38]Note that $\mathbf{D^*}$ is positive definite and symmetric in the case of quasilinear utility, as it is the negative of the Slutsky matrix, which is negative definite and symmetric. In that case, the solution is $\mathbf{b}^{GLS} = (\mathbf{F'D^*F})^{-1}\mathbf{F'D^*}\boldsymbol{\phi}$.

$$\frac{\partial \text{DWL}}{\partial \beta} = \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \times \sum_{s=1}^{S}(r_s - \beta) = 0. \tag{A.16}$$

Rearranging yields the solution: $\beta = \bar{r}$. Given this second-best policy, we now calculate the deadweight loss of the second-best policy and the deadweight loss of a constant unbiased tax. The deadweight loss from the second best policy is given by equation (A.15) for $\tau_j = \bar{r} f_j$:

$$\begin{aligned} -2\text{DWL}(\tau = \bar{r} f_j) &= \sum_{s=1}^{S}(r_s - \bar{r})^2 \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \\ &= S \times var(r_s) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}. \end{aligned} \tag{A.17}$$

The constant unbiased tax equals $\tau = k = \bar{r}\bar{f}$. The resulting deadweight loss is:

$$\begin{aligned} -2\text{DWL}(\tau = k) &= \sum_{s=1}^{S}\sum_{j=1}^{J}\sum_{k=1}^{J} e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \\ &= \sum_{s=1}^{S}\sum_{j=1}^{J}\sum_{k=1}^{J} (r_s f_j - \bar{r}\bar{f})(r_s f_k - \bar{r}\bar{f}) \frac{\partial x_{js}}{\partial t_{ks}} \\ &= \sum_{s=1}^{S} r_s^2 \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2\bar{f}^2 \times \sum_{j=1}^{J}\sum_{k=1}^{J} \frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2\bar{f} \times \sum_{j=1}^{J}\sum_{k=1}^{J}(f_j + f_k)\frac{\partial x_{js}}{\partial t_{ks}} \\ &= S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \end{aligned} \tag{A.18}$$

where, as detailed above in the derivation of equation A.15, the third equality follows from common markets (assumption 6) and the latter two terms are zero because of the no outside good assumption (assumption 7).

The fraction of the first-best welfare gain achieved by this second-best policy over a policy of a constant unbiased tax on all products can now be calculated as:

$$\begin{aligned} \frac{DWL(\tau = \bar{r} f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} &= \frac{\frac{1}{2} S \times (E(r_s^2) - var(r_s)) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{\frac{1}{2} S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}} \\ &= 1 - \frac{var(r_s)}{E[r_s^2]}. \end{aligned} \tag{A.19}$$

Proof of Corollary 2

This follows from the definition of R^2 of a regression of $\phi_{js} - r_s\bar{f}$ on f_j. The total sum of squares is:

$$TSS = \sum_{s=1}^{S}\sum_{j=1}^{J}(\phi_{js} - r_s\bar{f})^2 = \sum_{s=1}^{S}\sum_{j=1}^{J}(r_s(f_j - \bar{f}))^2 = \sum_{s=1}^{S} r_s^2 \sum_{j=1}^{J}(f_j - \bar{f})^2$$
$$= S \times E[r_s^2] \times \sum_{j=1}^{J}(f_j - \bar{f})^2. \tag{A.20}$$

A standard derivation of OLS shows that the slope is $\bar{r}$, and the constant is $-\bar{r}\bar{f}$. The OLS residuals from the regression are therefore given by $\phi_{js} - r_s\bar{f} - (\bar{r}f_j - \bar{r}\bar{f}) = r_s f_j - r_s\bar{f} - (\bar{r}f_j - \bar{r}\bar{f}) = (r_s - \bar{r})(f_j - \bar{f})$. Now compute the sum of squared residuals:

$$SSR = \sum_{s=1}^{S}\sum_{j=1}^{J}[(r_s - \bar{r})(f_j - \bar{f})]^2 = \sum_{s=1}^{S}(r_s - \bar{r})^2 \sum_{j=1}^{J}(f_j - \bar{f})^2$$
$$= S \times var(r_s) \times \sum_{j=1}^{J}(f_j - \bar{f})^2. \tag{A.21}$$

Now we can compute R^2 as:

$$R^2 = 1 - \frac{SSR}{TSS} = 1 - \frac{S \times var(r_s) \times \sum_{j=1}^{J}(f_j - \bar{f})^2}{S \times E[r_s^2] \times \sum_{j=1}^{J}(f_j - \bar{f})^2} = 1 - \frac{var(r_s)}{E[r_s^2]}. \tag{A.22}$$

Proof of Corollary 3

As demonstrated in equation (A.18), the deadweight loss for a constant unbiased tax $\tau = k = \bar{r}\bar{f}$ equals:

$$-2\text{DWL}(\tau = k) = S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2\bar{f}^2 \times \sum_{j=1}^{J}\sum_{k=1}^{J}\frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2\bar{f} \times \sum_{j=1}^{J}\sum_{k=1}^{J}(f_j + f_k)\frac{\partial x_{js}}{\partial t_{ks}}$$
$$= S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2\bar{f}^2 \times \sum_{j=1}^{J}\sum_{k=1}^{J}\frac{\partial x_{js}}{\partial t_{ks}} - 2S\bar{r}^2\bar{f} \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j \frac{\partial x_{js}}{\partial t_{ks}} \tag{A.23}$$

where the second and third terms do not cancel if assumption 7 does not hold and there is substitution to an outside good. The second equality follows from quasilinearity, which implies symmetry of the demand matrix.

Now define θ_j so that it solves, for each j, $-\theta_j \frac{\partial x_{js}}{\partial t_{js}} = \sum_{k \neq j} \frac{\partial x_{ks}}{\partial t_{js}}$. In words, θ_j is the total market size effect for a change in tax rate t_j; it is the ratio of the sum of cross-price effects (increases in quantity for other products that results from raising price j) to the own-price effect (decrease in quantity for product j from an increase in its price). If $\theta_j = 1$, there is no change in market size. If $\theta_j < 1$, the total market size (quantity summed across all J) shrinks as the price of j rises. Now rewrite expression (A.23):

$$\begin{aligned}
-2\text{DWL}(\tau = k) =& S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^{J} \left(\frac{\partial x_{js}}{\partial t_{js}} + \sum_{k \neq j} \frac{\partial x_{js}}{\partial t_{ks}} \right) \\
& - 2S\bar{r}^2 \bar{f} \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j \left(\frac{\partial x_{js}}{\partial t_{js}} + \sum_{k \neq j} \frac{\partial x_{js}}{\partial t_{ks}} \right) \\
=& S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^{J} (1 - \theta_j) \frac{\partial x_{js}}{\partial t_{js}} \\
& - 2S\bar{r}^2 \bar{f} \times \sum_{j=1}^{J} f_j (1 - \theta_j) \frac{\partial x_{js}}{\partial t_{js}} \\
=& S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2 \bar{f}^2 \times J(1 - \bar{\theta}) \overline{\frac{\partial x_j}{\partial t_j}} \qquad \text{(A.24)}
\end{aligned}$$

where the last equality follows from assumption 3 and two other simplyfying assumptions that we make here: f_j and θ_j are uncorrelated, and θ_j and $\frac{\partial x_j}{\partial t_j}$ are uncorrelated. These are assumptions for expositional convenience—they imply that the overall market size effects of different products are not related to the attribute or own-price derivatives and they make the result particularly clear, as the summation terms collapse to expressions of the mean.

In this case, the sign of the second term in equation (A.24) depends upon whether $\bar{\theta}$ is greater than, equal to, or less than 1. (Without the simplifying assumptions, the same sign is pivotal, but the result will depend on weighted averages of θ.) When $\bar{\theta} = 1$, the second term in (A.24) equals zero and Proposition 1 holds. The overall market size (outside good) effect grows in importance when $\bar{\theta}$ deviates more from 1, or when demand is more elastic ($\overline{\frac{\partial x_j}{\partial t_j}}$ is more negative). Note that is would be non-standard for $\bar{\theta} > 1$, which would imply a "sectoral Giffen good"—that is, on average across products in a sector, increase in individual product prices cause the overall market to expand. We thus assume that $\bar{\theta} \leq 1$.

We can now evaluate the relative welfare improvement of the second-best linear tax vs. the constant tax:

$$\frac{DWL(\tau = \bar{r}f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)}$$

$$= \frac{-\frac{1}{2}S \times var(r_s) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + \frac{1}{2}S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \frac{1}{2}JS\bar{r}^2\bar{f}^2(1-\bar{\theta})\overline{\frac{\partial x_j}{\partial t_j}}}{\frac{1}{2}S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \frac{1}{2}JS\bar{r}^2\bar{f}^2(1-\bar{\theta})\overline{\frac{\partial x_j}{\partial t_j}}}$$

$$= 1 - \frac{S \times var(r_s) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - JS\bar{r}^2\bar{f}^2(1-\bar{\theta})\overline{\frac{\partial x_j}{\partial t_j}}}. \tag{A.25}$$

It follows directly from equation (A.25) that $\bar{\theta} < 1$ implies that R^2 understates the fraction of the first-best welfare gain achieved by this second-best policy over a policy of a constant tax on all products:

$$\frac{DWL(\tau = \bar{r}f_j) - DWL(\tau = k)}{DWL(\tau = \phi) - DWL(\tau = k)} = 1 - \frac{S \times var(r_s) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - JS\bar{r}^2\bar{f}^2(1-\bar{\theta})\overline{\frac{\partial x_j}{\partial t_j}}}$$

$$> 1 - \frac{S \times var(r_s) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^{J}\sum_{k=1}^{J} f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}} = 1 - \frac{var(r_s)}{E[r_s^2]}. \tag{A.26}$$

B Appendix: Additional Empirical Results

Sensitivity to outliers

Table B.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.

Policies with a "slope bias"

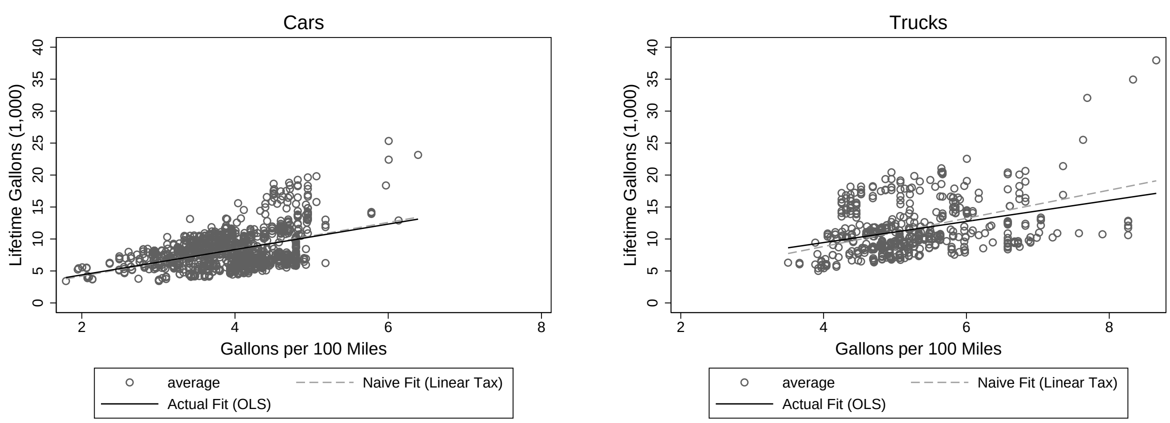
We also show here results related to the "naïve" linear tax, which is the OLS fit that would be chosen if there is no correlation between fuel economy and average lifetime mileage. We illustrate this in Figure B.1, which replicates Figure 3, but adds a line that represents the relationship between

Table B.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.

Figure B.1: The Relationship Between Lifetime Gasoline Consumption and Fuel-Efficiency



Note: The unit of observation is a type of vehicle (VIN10-prefix). 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.

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. Current fuel economy standards such as CAFE are naïve in this way, as the standards are not based on expected VMT. Figure B.1 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.