

# The Use of Regression Statistics to Analyze Imperfect Pricing Policies

Mark R. Jacobsen,<sup>1,5</sup> Christopher R. Knittel,<sup>2,5</sup> James M. Sallee,<sup>3,5</sup> Arthur A. van Benthem<sup>4,5\*</sup>

January 18, 2018

## Abstract

Corrective taxes can completely solve a variety of market failures, but actual policies are commonly forced to deviate from the theoretical ideal due to administrative or political constraints. This paper presents a method that requires a minimum of market information to quantify the efficiency costs of constraints on the design of externality-correcting tax schemes, or more generally the costs of imperfect pricing, using simple regression statistics. 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, has an immediate welfare interpretation—it characterizes the relative welfare gains achieved by alternative policies. We utilize our approach in four diverse empirical applications: random mismeasurement in externalities, imperfect spatial policy differentiation, imperfect electricity pricing, and heterogeneity in the longevity of energy-consuming durable goods. In two cases, we find that policy constraints are relatively harmless, while in the other two cases, the constraint induces large inefficiencies. Regarding the case of durable longevity, we find that policies that regulate vehicle fuel economy, but ignore the differences in average longevity across 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, second-best, externalities, imperfect pricing, energy efficiency  
JEL: H23, Q58, L51

---

\*A prior version of this paper was circulated under the title “Sufficient Statistics for Imperfect Externality-Correcting Policies.” We thank Eduardo Azevedo, Richard Blundell, Meghan Busse, Don Fullerton, Alex Gelber, Jeff Grogger, Jean-François Houde, Kelsey Jack, Damon Jones, Ravi Kanbur, Henrik Kleven, Olivia Mitchell, Juan-Pablo Montero, Joel Slemrod, Kent Smetters, Johannes Spinnewijn and seminar participants at the 2015 AEA meetings, 2015 AERE conference, Spring 2015 NBER EEE meetings, 2015 Northeast Workshop on Energy Policy and Environmental Economics, 2015 NTA conference, Catholic University of Chile, Colorado School of Mines, Cornell, Free University Amsterdam, Imperial College, INSEAD, Institute for Fiscal Studies, London School of Economics, Michigan State University, NYU Abu Dhabi, Stanford University, Toulouse School of Economics, University College London, UC Berkeley, UCLA, University of Cambridge, University of Chicago, University of Chile, University of Colorado Boulder, University of Oxford and the University of Virginia for helpful comments and suggestions. Van Benthem thanks the National Science Foundation (award number SES1530494), the Kleinman Center for Energy Policy at the University of Pennsylvania, and the Wharton Dean’s Research Fund for support. Sallee thanks Google for support through a gift to the Energy Institute at Haas.

<sup>1</sup>Department of Economics, University of California San Diego. <sup>2</sup>Sloan School of Management and Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology. <sup>3</sup>Department of Agricultural and Resource Economics, University of California, Berkeley. <sup>4</sup>The Wharton School, University of Pennsylvania. <sup>5</sup>National Bureau of Economic Research.

# 1 Introduction

Many important public policies aim to fix market failures that create wedges between marginal social costs and benefits. Many prominent examples are externality-correcting policies, which range from taxes on cigarettes, alcohol, or sugary beverages to mandatory immunizations to the regulation of pollution. Since [Pigou \(1932\)](#), economists have understood that, if there are no additional market failures beyond the externality, market efficiency can be fully restored when externalities are taxed directly and the marginal damage at the optimal quantity is known. Yet, relatively few policies closely follow this prescription. Often it is administratively impossible, technologically too costly, or politically infeasible to price actions according to the externalities that they generate.

Consequently, externality-correcting policies are generally imperfect. Imperfection often takes the following form: the externality is dependent on a set of variables, but policy is contingent on only a subset of those variables or their imperfect proxies. For example, the external damages from sulfur dioxide depend on the amount of pollution emitted, the weather, and the location of emissions relative to population centers. But, sulfur dioxide regulations are based only on emissions quantities. In transportation, congestion externalities are highly concentrated in certain times of day, but most toll prices are uniform or vary only slightly with traffic conditions. In health, the externalities associated with second-hand smoke depend on many factors, including proximity to other people, whether the smoking is indoors or outdoors, etc. But, cigarette taxes are uniform.

In this paper, we develop a model that characterizes the welfare costs of using policies that take this form. We show that, when certain conditions are met, familiar statistics from simple regressions of the true externality on the variables upon which policy is based have direct welfare interpretations. Specifically, deadweight loss scales with the sum of squared residuals, and the  $R^2$  summarizes the fraction of the welfare gain from a Pigouvian benchmark that is achievable by the constrained (which we call second-best) policy. We demonstrate the usefulness of the method through four empirical applications.

Our theory posits a standard model of a competitive market with a representative consumer who chooses among a variety of related goods, each of which produce a different level of an externality. A vector of Pigouvian taxes (which we call the Pigouvian benchmark) on these goods can restore efficiency, but we suppose that the planner faces a constraint, so that taxes must be made contingent upon some variable that is imperfectly correlated with the externality. This induces “errors” in the constrained optimal tax rates, as compared to the Pigouvian benchmark. We build on [Harberger \(1964\)](#) in deriving a general 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-price derivatives of demand, which will typically be unavailable. However, under some conditions regarding the demand matrix, second-best policies will involve a set of taxes or shadow prices under which cross-product substitution does not affect overall welfare. Intuitively, what is required is that two products that are closer substitutes for each other do not, on average, have more similar tax rate errors.

We show that, when this condition is met, welfare conclusions can be drawn with limited in-

formation. Given data on the distribution of the externality and its degree of correlation with the variables upon which policy is based, one can determine the *proportion* of the welfare gain achievable by the Pigouvian policy that the second-best policy achieves. Unlike results in the previous literature, this policy comparison does not require an estimate of any behavioral parameters. Given an estimate of the own-price derivative for the goods and the marginal damage due to the externality, the welfare costs of employing second-best policies in lieu of the Pigouvian benchmark can be estimated directly in dollars (rather than as a proportion).

To demonstrate the power of this method, we apply it to four distinct empirical problems. The first application considers random mismeasurement—energy efficiency is measured according to laboratory test procedures which differ from in-use averages, thereby creating mismeasurement in externalities 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 ratings. We conclude that the second-best policy is quite efficient; it obtains more than 95% of the gains achieved by the Pigouvian benchmark.

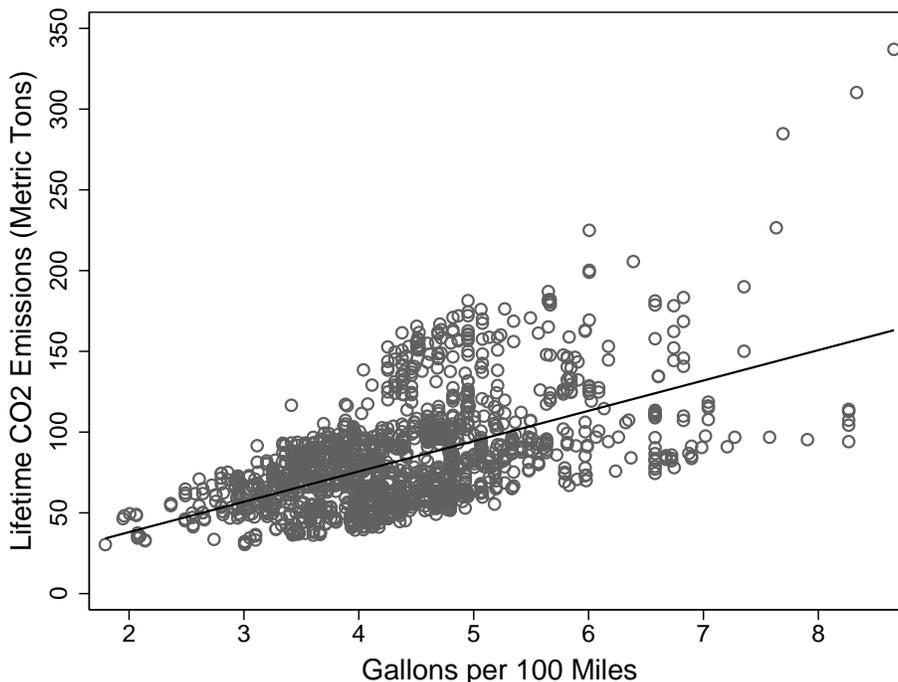
Our second application regards real-time electricity pricing. Unlike our other three applications, this does not concern an externality. Instead, there is a wedge between marginal costs and benefits due to the fact that the marginal cost of generating electricity varies hour by hour, but electricity tariffs do not vary to reflect these costs (Borenstein and Holland 2005). We apply our method to characterize the welfare gain of tariffs that vary along some time or date dimensions, but fall short of the theoretical ideal of real-time pricing. We find that realistic time-varying tariffs recover only a modest fraction of the gains achieved by real-time pricing.

Our 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 on 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. We quantify that average lifetime miles traveled by individual vehicles of a particular model vary substantially 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. We conclude that actual fuel-economy policies, which treat such vehicles identically, recover only about one-quarter to one-third of the welfare gain compared to a policy that considers both fuel economy and vehicle longevity. This result is robust even when we relax key assumptions about demand.

To illustrate our results, Figure 1 shows the relationship between fuel-economy ratings and average lifetime carbon emissions for different types of automobiles. Each data point represents the average lifetime CO<sub>2</sub> emissions across a number of individual vehicles of the same model (e.g., all 2012 Toyota Camry LE observations are combined into one data point). The solid line is the linear best fit. Dispersion in the data comes from heterogeneity in lifetime mileage; if all vehicles had the same lifetime mileage, the data would lie on a straight line. Federal fuel-economy standards impose implicit taxes on vehicles that are a linear function of each vehicle’s official fuel-consumption rating;

they cannot be based on average lifetime mileage. Our theory shows that, under some conditions, the second-best fuel-economy standard creates implicit taxes equal to the OLS prediction line and the  $R^2$  from this regression—0.29 in the case of Figure 1—is an estimate of the fraction of the Pigouvian welfare gain that is achieved by this fuel-economy policy.

**Figure 1:** The Relationship Between Lifetime CO<sub>2</sub> Emissions and Fuel Efficiency



Note: An observation is the average lifetime miles driven for a particular vehicle model, taken across many individual units, divided by the fuel-economy rating and multiplied by the tons of CO<sub>2</sub> 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 5. The solid line is an OLS regression line.

A fourth application considers spatial differentiation. A given amount of pollution or energy use may have quite different health or environmental consequences depending on where it takes place, but policies often cannot differentiate their treatment by location. We use our framework to quantify the welfare costs of imperfect spatial differentiation for the case of carbon dioxide emissions resulting from the use of electric appliances. Here, differences in emissions across space are due to the fact that the emissions rate from the marginal power plant differs across regions of the country. This application also serves to demonstrate the broader applicability of regression statistics for welfare analysis because our required demand conditions will not hold for the policy we consider. Instead, we demonstrate that an alternative regression statistic, the within- $R^2$  from a regression with spatial fixed effects, has the desired interpretation. We conclude, for this particular case, that the welfare costs of failing to spatially differentiate are small.

Our method in general, and our analysis of fuel-economy policy specifically, represent contribu-

tions to the evaluation of energy efficiency policies. No prior research has analyzed the implications of heterogeneity in lifetime utilization for the design of energy efficiency programs. This adds a new, and apparently economically important, dimension to the analysis of energy efficiency programs. In particular, it points out a new concern for the comparison between gasoline taxation and fuel-economy standards as competing policies aimed to reducing greenhouse gas emissions from transportation.<sup>1</sup> Similar issues arise for any policy that regulates pollution-control technology.

More broadly, our main contribution is to show the relationship between familiar regression statistics and second-best policies that aim to fix market failures but are constrained to be imperfect. This relates to the sufficient statistics literature in public finance, which is similar in seeking to find ways of characterizing welfare effects of policies that require information about a minimum number of parameters. Our analysis is unique in focusing on regression statistics, and also adds to the small set of articles in this literature that are focused on externalities.<sup>2</sup>

Our analysis also connects to an important strand of literature in environmental economics that considers heterogeneity in damages from the same pollutant emitted in different locations. For example, the marginal damage from a ton of sulfur dioxide will differ depending on whether or not it is emitted near a densely populated city. A theoretical literature has noted that this type of spatial heterogeneity implies that uniform national policies are inefficient, and suggested an efficiency gain from spatially differentiated regulation (Tietenberg 1980; Mendelsohn 1986; Baumol and Oates 1988). This type of concern has been used to study the potential benefits of spatial differentiation in policies regarding air pollution (Muller and Mendelsohn 2009; Muller, Mendelsohn, and Nordhaus 2011; Fowlie and Muller 2017), renewable energy generation (Cullen 2013; Callaway, Fowlie, and McCormick 2018), water pollution (Farrow, Schultz, Celikkol, and van Houten 2005), and electric vehicles (Holland, Mansur, Muller, and Yates 2016). As we discuss in Section 6, a number of these models can be understood as special cases of our general setup, and we suggest that our approach could offer a straightforward way of estimating potential gains from counterfactual policies in these contexts.

---

<sup>1</sup>For reviews of this literature for automobiles, see Harrington, Parry, and Walls (2007); Anderson, Parry, Sallee, and Fischer (2011); Anderson and Sallee (2016). Existing research, including Fullerton and West (2002), Fullerton and West (2010) 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, none consider heterogeneity in average lifetime utilization.

<sup>2</sup>Chetty (2009) documents a broad set of topics that have been considered by the literature on sufficient statistics in public economics, but he cites no papers focused on externalities. Recent work has included not only traditional questions in taxation (Feldstein 1999; Goulder and Williams 2003; Kleven and Kreiner 2006; Saez, Slemrod, and Giertz 2012; Hendren 2016), but also studies of social insurance (Baily 1978; Chetty 2006), health insurance (Einav, Finkelstein, and Cullen 2010), and limited rationality (Chetty, Looney, and Kroft 2009; Allcott, Mullainathan, and Taubinsky 2014). Hendren (2016) briefly notes that, in order to fully assess a policy in the presence of externalities, one needs to know the effect of the policy on the externality net of many general equilibrium (cross-price) effects across a variety of related goods. But that paper does not propose a way to estimate this net effect, whereas we described conditions when they will cancel. One paper that invokes the sufficient statistics tradition and does explicitly consider energy is Allcott, Mullainathan, and Taubinsky (2014), which models energy efficiency policy when heterogeneous consumers may undervalue energy efficiency due to limited rationality. They model a discrete choice between an efficient or inefficient good and derive sufficient statistics for the optimal combination of energy taxes and subsidies for energy efficient products.

We investigate four distinct empirical applications in this paper, and we believe that the methods can be applied more broadly. The key data requirement is some measure of the distribution of the externality (or other efficiency wedge) and its correlation with the variables upon which policy is contingent. For a few more examples, consider the policies we described at the beginning of this introduction. The efficiency of sulfur dioxide trading programs could be assessed using estimates of the spatial distribution of marginal damages generated by [Muller and Mendelsohn \(2009\)](#) and [Muller, Mendelsohn, and Nordhaus \(2011\)](#).<sup>3</sup> The efficiency of various congestion pricing policies could be estimated using existing traffic data, such as the high frequency records from thousands of locations in the California highway system ([Caltrans 2016](#)). Data on second-hand smoke exposure at home and in the workplace from the National Adult Tobacco Survey for the U.S. or the Global Adult Tobacco Survey could be used to estimate the efficiency of cigarette taxes as tools for mitigating externalities from second-hand smoke.<sup>4</sup>

The balance of the paper is as follows. In [Section 2](#) we develop the theory for deriving sufficient statistics. In [Section 3](#) we apply our method to the case of random mismeasurement in externalities, using a recent change in fuel-economy testing procedures for automobiles. [Section 4](#) shows how our method applies to mispricing in electricity markets. In [Section 5](#) we apply our results to heterogeneity in the longevity of automobiles. [Section 6](#) considers spatial heterogeneity in emissions from identical products used in different locations, using carbon emissions from refrigerators as an example. [Section 7](#) concludes.

## 2 Theory for Deriving Sufficient Statistics

The goal of our model is to facilitate analysis of the efficiency costs of policies that correct an externality or another wedge between marginal costs and benefits but that deviate from the theoretical ideal. Actual policies may be less efficient than an ideal policy for a variety of reasons, including political constraints, technological cost, and administrative feasibility. After presenting our model setup and notation, we first derive a general expression for the welfare loss from using some alternative, constrained policy in lieu of the ideal. We then specify sufficient conditions under which this general expression collapses so that simple regression statistics have welfare interpretations. Finally, we describe what can be learned from simple regression statistics even when our sufficient conditions are not met and when externalities are measured with error.

### 2.1 Model setup

We emphasize a simple model in which there is only one market failure. We model a representative consumer in a perfectly competitive market. The economy has products indexed  $j = 1, \dots, J$ .

---

<sup>3</sup>Spatial heterogeneity is not the only factor that determines the efficiency of SO<sub>2</sub> trading. [Montero \(1999\)](#), for example, demonstrates that adverse selection in voluntary opt-in to the SO<sub>2</sub> trading program in the U.S. had significant efficiency impacts in the program's early years.

<sup>4</sup>For details on those data sources, see [http://www.cdc.gov/tobacco/data\\_statistics/surveys/nats/index.htm](http://www.cdc.gov/tobacco/data_statistics/surveys/nats/index.htm) and <http://www.who.int/tobacco/surveillance/gats/en/>.

The consumer chooses quantities of each, denoted  $x_j$ . The consumer derives utility,  $U$ , from the consumption of these products according to the function  $U(x_1, \dots, x_J)$ , which we assume is twice differentiable, increasing, and weakly concave in each argument. We denote the cost of production by  $C(x_1, \dots, x_J)$ , 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 quasilinear numeraire,  $n$ . We assume no technological change and do not model the endogenous entry and exit of products into the market.<sup>5</sup> As such, ours is a short-run model, though one could allow for zero quantities so that the product vector represents potential products.

We posit that there is some market failure which leads the market, absent policy, to choose quantities so that there is a wedge, denoted  $\phi_j$ , between the marginal private benefit and the marginal private cost of a unit of  $x_j$ . Our first assumption is that  $\phi_j$  is fixed and unchanging with respect to policy intervention, and that the total social inefficiency is the sum of these wedges across goods, multiplied by quantities:  $\phi = \sum_{j=1}^J \phi_j x_j$ . The simplest interpretation is that  $\phi_j$  is an externality, as is the case in three of our four applications. In one of our applications,  $\phi_j$  is a gap between marginal cost and marginal benefit due to coarse pricing, where the cost of producing a good varies over time but the price is constrained to be constant.<sup>6</sup> Inspired by the externality interpretation, we refer to  $\phi_j$  as the marginal social damage per unit of  $x_j$ .

**Assumption 1.** *Marginal social damages from each product,  $\phi_j$ , are fixed w.r.t. the tax vector  $t$ .*

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,  $\phi_j$ , 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  $t_j$ . We describe policies as taxes on products, but this is equivalent to regulatory policies that create implicit taxes (shadow prices). 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. The consumer’s optimization problem is:

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

---

<sup>5</sup>For a treatment of how product redesigns can influence the design of a tax system that is limited in its ability to assign unique tax rates to each product, see [Gillitzer, Kleven, and Slemrod \(2017\)](#).

<sup>6</sup>The wedge could come from other sources, such as market power, but our derivation assumes that  $\phi_j$  is fixed with respect to the policy vector. Markups will generally shift with policy intervention, so application of our framework to market power would require modifications.

The consumer's first-order conditions imply that  $\frac{\partial U}{\partial x_j} = p_j + t_j$ , which we assume holds at an interior solution.

Social welfare  $W$  is the utility from the product bundle, the numeraire (substituted out for the budget constraint), and the externality:

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

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:

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

The first-order condition for product  $j$  is:

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

where the second equality follows from substituting the consumer's first-order condition, and from our assumption of marginal cost pricing.

For wedges other than an externality, the same expression will arise as long as the wedge satisfies Assumption 1. For example, one of our applications relates to coarse pricing—electricity prices are constant at all hours of the day, whereas marginal cost varies. In this case,  $\phi_j$  is the gap between the price faced by the consumer (which equals marginal utility) and the true marginal cost. To consider that case, drop the externality term from Equation 2. Then, differentiating  $W$  with respect to  $t_j$  yields the same  $\sum_{k=1}^J (t_k - \phi_k) \frac{\partial x_k}{\partial t_j}$ .

Equation 4 shows that 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 optimum is a vector of Pigouvian taxes; each product's tax rate is set equal to its marginal external damage. This is as expected. We refer to the policy vector  $t_j = \phi_j \forall j$  as the Pigouvian benchmark.

We wish to characterize how welfare under this Pigouvian benchmark compared to that under a policy that satisfies some constraint. The difference represents the cost of the constraint on policy design. The constraint is a restriction on the vector of taxes that the planner can choose. We write this constraint as a function  $t_j = g(f_j; \theta)$ , where  $f$  is some vector of exogenous attributes of the products,  $\theta$  are parameters to be chosen by the planner, and  $g$  is some function. The planner's problem can now be written:

$$\begin{aligned} \max_{\theta} W &= U(x_1, \dots, x_J) + M - C(x_1, \dots, x_J) - \sum_{j=1}^J \phi_j x_j \\ \text{s.t. } t_j &= g(f_j; \theta) \quad \forall j \end{aligned} \tag{5}$$

We call the solution to this policy, denoted  $t_j = g(f_j; \theta^*)$ , the second-best, or constrained, tax vector. Recall that our goal is to provide welfare interpretations of regression statistics. Motivated by this, we restrict attention to situations where  $g(f_j; \theta)$  can be written as linear in parameters, noting that this is no more restrictive than it is in any application of (multivariate) linear regression, where variables can be transformed and interacted. For example, in our third application, we consider fuel-economy regulations that impose a shadow tax on vehicles that is an affine function of their fuel-economy ratings. Thus,  $g(f_j; \theta) = t_j = \alpha + \beta f_j$ , where  $\theta$  consists of two parameters,  $\alpha$  and  $\beta$ , and  $f_j$  is the fuel-economy rating. Our four applications demonstrate a variety of policy-design constraints that fit into this framework.

Our objective is to describe the welfare cost of such policy constraints relative to the Pigouvian benchmark. We note that the Pigouvian benchmark itself is not necessarily “first-best” in the presence of other market failures or margins of adjustment that product-based taxes cannot correct.<sup>7</sup> For example, taxes on new vehicles cannot induce optimal scrappage behavior. Therefore, our Pigouvian new vehicle tax vector falls short of a first-best tax on gasoline. We discuss this in more detail in Section 5. In such cases, our method considers the welfare gain along a particular dimension of interest that is targeted directly by the tax, assuming that other distortions are held constant. We return to this point in the applications.

To describe the welfare consequences of such policy constraints, we now proceed to deriving a generic expression that characterizes the loss of social welfare caused by moving from the Pigouvian benchmark policy to some arbitrary tax vector. We then use this expression to relate second-best policies that would arise given a particular constraint.

## 2.2 Characterizing deadweight loss

Let a generic tax schedule be denoted as  $\tau_1, \dots, \tau_J$ . We characterize the welfare loss of moving from the Pigouvian benchmark  $t_j = \phi_j$  to  $t_j = \tau_j$  by specifying a weighted average of the two tax schedules and then integrating the marginal welfare losses of moving the weights from  $\phi_j$  to  $\tau_j$ . We denote the difference in welfare between the two schedules as  $DWL(t = \tau)$ .<sup>8</sup> To do so, we assume local linearity.

---

<sup>7</sup>For simplicity, even though the Pigouvian benchmark will not be first-best in all settings, we refer to the constrained policy as “second-best”.

<sup>8</sup>We consider relative benefits of policies but the final policy choice also depends on relative costs (e.g., administrative or technology costs). Our method provides a bound on the costs that would make the less precise policy better overall. Note that, in our four empirical examples, the Pigouvian benchmark is technically feasible and requires only a set of tax rates or prices, where taxes and prices are already being charged. It seems unlikely then that administrative costs would be a major factor in comparing policies in our settings, but this may not be true in other situations.

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

Under the assumption of constant demand derivatives, the efficiency loss incurred from imposing any arbitrary tax schedule  $\tau$  in lieu of the Pigouvian tax schedule can be written as:

$$W(t = \phi) - W(t = \tau) \equiv DWL(t = \tau) = -\frac{1}{2} \sum_{j=1}^J \sum_{k=1}^J (\tau_j - \phi_j) (\tau_k - \phi_k) \frac{\partial x_j}{\partial t_k}. \quad (6)$$

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 (although without externalities) is in Harberger (1964). When  $\tau_j = \phi_j$ , each term in the summation will be zero.

In line with the traditional use of Harberger triangles, we assume that demand derivatives are constant over the relevant range of taxes. In our discussion here, we also assume that producer prices are unchanged, which implies constant marginal cost. In this case,  $\frac{\partial x_j}{\partial t_k}$  represents only a demand derivative, not a combined effect of supply and demand. Where marginal cost is increasing but linear, our mathematical results are all the same, but  $\frac{\partial x_j}{\partial t_k}$  is interpreted as the combined response of supply and demand (see Appendix A).

Where demand or marginal costs are convex, our results represent a local approximation in the same way that Harberger triangles normally do. Thus, our derivations can also be understood as indicating incremental welfare losses from small movements away from the Pigouvian benchmark. Note that we relax both the linearity and the constant marginal cost assumptions in our electricity-pricing application, but we preserve them here for the exposition.

To better understand the content of Equation 6 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 effects:

$$-2 \times DWL(t = \tau) = \sum_{j=1}^J \sum_{k=1}^J e_j e_k \frac{\partial x_j}{\partial t_k} \quad (7)$$

$$= \underbrace{\sum_{j=1}^J e_j^2 \frac{\partial x_j}{\partial t_j}}_{\text{own effects}} + \underbrace{\sum_{j=1}^J \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}_{\text{cross effects}}. \quad (8)$$

(7) and (8) are quite general expressions. But, using these formulas to evaluate 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.<sup>9</sup> Specifically, the cross effects in Equation 8 will be zero when there is

<sup>9</sup>Goulder and Williams (2003) also build from the general Harberger formula and present a simplified expression for the excess burden of taxation that does not require estimates of all cross-price derivatives. They study interactions between commodity and labor taxes, a very different setting from ours.

no substitution between goods, so that cross-price derivatives are all zero. Alternatively, the cross effects will be proportional to the own effects when the errors in the tax rates are mean zero and products of errors are uncorrelated with cross-price derivatives. Note that we refer to  $\partial x_j / \partial t_k$  as cross-price *derivatives* and the contribution of the  $e_j e_k \partial x_j / \partial t_k$  to deadweight loss as cross *effects*. Zero cross-price derivatives are a sufficient condition for the cross effects to simplify, but so are the alternative conditions listed below.

We state these possibilities formally as Assumption 3. We then proceed to derive results under the case where Assumption 3 holds before returning to a detailed discussion of these conditions.

**Assumption 3.** (a) Tax errors  $e_j$  are uncorrelated with own-price derivatives:  $\text{cov}\left(e_j, \frac{\partial x_j}{\partial t_j}\right) = 0$ .  
 (b) Products of tax errors  $e_j e_k$  are uncorrelated with cross-price derivatives:  $\text{cov}\left(e_j e_k, \frac{\partial x_j}{\partial t_k}\right) = 0 \forall j \neq k$ . (A stronger version of this, (b'), assumes that cross-price derivatives are zero:  $\frac{\partial x_j}{\partial t_k} = 0 \forall j \neq k$ .)

where  $\text{cov}\left(e_j e_k, \frac{\partial x_j}{\partial t_k}\right)$  is calculated for all non-diagonal elements of the demand matrix ( $j \neq k$ ). Version (b') of this assumption holds if there is no substitution across products. Version (b) assumes that cross-price derivatives between each pair of products are uncorrelated with the product of their tax errors. This holds if externalities, conditional on policy, are orthogonal to substitutability. As we discuss further below, this is a plausible property of second-best policies. To provide a more intuitive economic interpretation, we note that one way that this assumption can be satisfied is if i) for each product  $j$  the errors of its substitutes are uncorrelated with the cross-price derivatives and ii) across products  $j$  the tax errors are uncorrelated with average cross-price derivatives.

In our empirical applications, we provide examples where (b') is likely to hold by approximation (electricity pricing) as well as as cases in which (b) is reasonable (fuel-economy standards and noisy energy efficiency ratings). Nevertheless, Assumption 3 will not hold in all cases, so we review what can be learned when the conditions do not hold after establishing our primary results that obtain under Assumption 3. Moreover, we provide numerous robustness checks throughout our empirical applications.

### 2.3 Welfare statistics when DWL is proportional to squared tax errors

Under Assumptions 1, 2, the strong version of Assumption 3 (parts (a) and (b')), and assuming unbiasedness on average so that  $\sum_{j=1}^J e_j = 0$ , the deadweight loss of an arbitrary tax vector is given by:

$$\text{DWL} = -\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} \sum_{j=1}^J e_j^2, \quad (9)$$

which foreshadows a central role for minimizing a sum of squared tax errors. When we use the weaker variant of Assumption 3 (parts (a) and (b)), the welfare loss expression remains very similar:

$$DWL = -\frac{1}{2} \left( \overline{\frac{\partial x_j}{\partial t_j}} - \overline{\frac{\partial x_j}{\partial t_k}} \right) \sum_{j=1}^J e_j^2. \quad (10)$$

where the average cross-price derivative (over the non-diagonal entries of the demand derivative matrix)  $\overline{\frac{\partial x_j}{\partial t_k}} = \frac{1}{J(J-1)} \left( \sum_{j=1}^J \sum_{k \neq j} \frac{\partial x_j}{\partial t_k} \right)$ . Note that DWL in Equation 10 is still proportional to the sum of squared tax errors, but it is multiplied by the difference in the average own-price and the average cross-price derivative. Because the proportionality of the DWL is maintained, all propositions and corollaries below hold exactly using either variant of Assumption 3. The algebra leading to Equations 9 and 10 appears in Appendix A.

When the number of goods  $J$  is large,  $\overline{\frac{\partial x_j}{\partial t_k}}$  will become small. Thus, Equation 9 will be a close approximation of the DWL in Equation 10 even in cases where only the average own-price elasticity is known. Moreover, even for smaller values of  $J$ ,  $\overline{\frac{\partial x_j}{\partial t_k}}$  will shrink if the substitution to the outside good becomes larger.

The solution to the planner's constrained problem in Equation 5 is the same as from minimizing the deadweight loss in Equation 10 subject to the same constraint.<sup>10</sup> This makes the link between policy and regression obvious. Whenever the policy constraint  $t_j = g(f_j; \theta)$  can be written as a function that is linear in parameters, minimization of deadweight loss is the same as minimizing the sum of squared residuals in a regression of the true externalities on the tax rates.

When Assumptions 1 to 3 hold, the second-best policy will be to choose  $\alpha$  and  $\beta$  to be the OLS solutions from fitting the externality to the policy variable. This is stated in Proposition 1:<sup>11</sup>

**Proposition 1.** *Under Assumptions 1 to 3, the second-best policy is the OLS fit of  $\phi_j$  to  $f_j$ , and the deadweight loss is proportional to the sum of squared residuals:*

$$DWL = -\frac{1}{2} \left( \overline{\frac{\partial x_j}{\partial t_j}} - \overline{\frac{\partial x_j}{\partial t_k}} \right) SSR. \quad (11)$$

The proof is in Appendix A. The intuition is as follows. When the externalities, conditional on characteristics that are in the policy function, are uncorrelated with product substitutability then the deadweight loss is a linear function of the sum of squared tax errors and the sum of errors (bias in the tax) squared. We show that this objective function is minimized by the same line that minimizes the sum of squared tax errors: a simple OLS fit. We show below how weighted least squares provides a similar solution when own-price derivatives may be correlated with the error.

In turn, the resulting deadweight loss is the sum of squared residuals from the OLS regression scaled by the average demand derivative and an average cross-price derivative factor that is close

<sup>10</sup>Deadweight loss is just the objective function evaluated at the Pigouvian benchmark minus the objective function evaluated at an alternative tax vector, so the original objective function is just deadweight loss plus a constant term.

<sup>11</sup>For expositional ease, we derive results for the case where policy is contingent on one 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  $\alpha$  and  $\beta$  where  $t_j = \alpha + \beta f_j$ . It is straightforward to modify our derivation to include many variables.

to zero when the number of products  $J$  is large. Thus, given data on the externality and own-price derivatives, and the attributes upon which policy is based, an analyst can run a simple linear regression and assign direct welfare interpretations to the regression output.

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 uniform unbiased tax rate  $\bar{t}$  that is the same for all products.<sup>12</sup>

**Corollary 1.** *Under Assumptions 1 to 3, the  $R^2$  from the OLS fit of  $\phi_j$  to  $f_j$  represents the percentage of the welfare gain of the Pigouvian tax (relative to a baseline of a uniform unbiased tax  $\bar{t}$ ) that is achieved by the second-best linear tax on  $f_j$  (relative to the same baseline):*

$$R^2 = \frac{DWL(t = \alpha^{OLS} + \beta^{OLS} f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})}. \quad (12)$$

Under our assumptions, the  $R^2$  relaxes the information requirement of knowing own-price derivatives and also eliminates the small adjustment factor involving the average cross-price derivative. No moments of the demand system are required to calculate this sufficient statistic. 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.

We now relax part (a) of Assumption 3 to allow correlation between errors and own-price derivatives. This leads to a very intuitive relationship with weighted multivariate regression:

**Proposition 2.** *Under Assumptions 1, 2 and 3(b), the second-best policy is the weighted least squares fit of  $\phi_j$  to a vector of attributes  $f_j$ , where the weighting matrix is diagonal with each entry equal to the own-price derivative for product  $j$ .*

Given information about the own-price derivatives of each product, a researcher could calculate the WLS estimator and derive parallel welfare results for this case. The second-best policy is still a linear best fit; the deadweight loss is the weighted sum of squared residuals from that regression. The proof appears in Appendix A. Further, when relaxing Assumption 3 altogether, the second-best policy is the GLS fit of  $\phi$  to  $f$  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 this formula would be useful for robustness analysis. We demonstrate such a calculation in Section 5.6, using estimates on the full matrix of demand elasticities for automobiles.

---

<sup>12</sup>The application of the deadweight loss formula in Equation 10 to this benchmark requires applying Assumptions 1 to 3. When we relax Assumption 3 in Section 2.4 we also relax its application to the benchmark policy.

## Interpreting the assumptions about cross effects

We now discuss the key economic implications of the assumptions needed for the results above. Part (a) of Assumption 3 says that the strength of own-price derivatives is not correlated with a product’s tax error; i.e., whatever factors that determine the externality but are omitted from the policy function do not also indicate stronger or weaker own-price responses. Proposition 2 relaxes this assumption. Doing so is important in empirical applications where some products are demanded in much larger quantities than others (and so have larger own-price derivatives, all else equal).

The second part of Assumption 3 has more economic content. The strong version 3(b’) applies to markets where products are not substitutes or complements. This assumption is unlikely to hold, though we argue in Section 4 that it applies, at least by approximation, to electricity pricing. Even when cross-price derivatives are not zero, Corollary 1 will still apply as long as the weaker version of 3(b) holds: this says that the difference between the errors in the tax rates between two products 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 the attributes upon which policy is contingent,  $f$ . Consider the vehicle example. Two vehicles with similar externalities ( $\phi$ ) will be closer substitutes, provided that vehicle fuel economy ( $f$ ) is a factor that determines vehicle choice, because  $\phi$  is mechanically related to  $f$ . But, Assumption 3 can still be met if, after conditioning on fuel economy, the residual variation in the externality  $\phi$  is not correlated with substitutability. Whether this will be true depends on the variables that are included in the policy and the source of residual variation in the externality. We discuss Assumption 3 in more detail for each of our empirical applications.

The results in this section demonstrate that—under assumptions that are often plausible—the deadweight loss of deviating from the Pigouvian benchmark can be calculated with limited information about the market. The welfare gains possible in the second-best relative to those in the Pigouvian case can be calculated with even less information. In the next four sections we demonstrate that these theoretical results have empirical relevance by illustrating four situations in which a sufficient statistic useful for evaluating policy can be derived from this framework.

### 2.4 What information remains in the $R^2$ when the cross effects do not simplify?

In this subsection, we explore cases where the  $R^2$  is biased (because our assumptions do not hold), but that bias can be signed, so the  $R^2$  is interpretable as a bound on welfare effects. To be precise, we consider what the  $R^2$  indicates about the welfare gains from the linear best fit policy, showing when this overstates or understates welfare gains. When our assumptions do not hold, this linear best fit policy may not be second-best. But, we still think it is the most interesting candidate policy to analyze for many situations where the policy-maker lacks the detailed information about demand needed to determine how the second-best deviates from the best fit. We first present a formula that highlights how different forces push the true welfare ratio away from  $R^2$  in different directions. We then make suggestions for how empiricists might investigate the potential bias based

on the type of violation.<sup>13</sup>

When we do not impose Assumption 3 so that cross effects do not simplify, we can still write out an expression for the relative gain in welfare achieved by the linear-best fit policy over a uniform tax policy, divided by the gain from the Pigouvian benchmark over the same uniform tax. We denote this welfare gain by  $S$ , and compare it to the  $R^2$ :

$$S = 1 - \frac{-\frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} SSR^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}{-\frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} TSS^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} \lambda_j \lambda_k \frac{\partial x_j}{\partial t_k}}, \quad (13)$$

where  $\lambda_j$  are the residuals in the regression of  $\phi$  on a constant (the uniform policy). Note that  $\lambda_j = \gamma_j + e_j$ , where  $\gamma_j$  is defined as the explained portion in the linear regression:  $\gamma_j = \alpha^{OLS} + \beta^{OLS} f_j - \bar{\phi}$ . Because  $\gamma$  is a function of  $f$ , the tax errors in the uniform policy depend on  $f$ . Thus, Equation 13 allows cross-price derivatives to be correlated either with  $e$  or with  $f$  (and thus  $\lambda$ ).

We now consider two types of correlation that determine the direction of the bias in  $R^2$ . First, correlation between products of constrained policy errors and cross-price derivatives  $cov\left(e_j e_k, \frac{\partial x_j}{\partial t_k}\right)$  (“type 1”). Second, correlation between products of the policy variable and cross-price derivatives  $cov\left(\gamma_j \gamma_k, \frac{\partial x_j}{\partial t_k}\right)$  (“type 2”).

**Proposition 3.** *Under Assumptions 1 and 2, (i)  $R^2 < S$  if type 1 correlation is positive (but type 2 correlation is zero); (ii)  $R^2 > S$  if type 2 correlation is positive (but type 1 correlation is zero).*

The proof is in Appendix A. First consider part (i) of Proposition 3. If cross-price elasticities are larger for goods with similar tax errors (e.g., vehicle durability in application 3) then the true fraction of welfare recovered in the second-best policy increases relative to the  $R^2$  measure. The intuition here is that when goods with similar tax errors are good substitutes, the Pigouvian benchmark loses some of its advantage: consumers do not substitute much along this dimension anymore and so the two policies become more similar, acting mostly along the margin of reducing  $f$ .

Now consider part (ii). If cross-derivatives are large when  $\lambda_j$  and  $\lambda_k$  (a function of the observable attribute  $f$ , such as fuel economy in application 3) are similar then the true fraction of welfare recovered by the second-best policy  $S$  will decrease relative to  $R^2$ . The intuition for this follows from observing that correlation of substitutability with  $f$  makes the second-best policy less effective because consumers now substitute mainly among products with similar  $f$ . The Pigouvian benchmark is still based on both  $f$  and the tax error and so its effectiveness is not damaged as much.

It is important to note that, in many common settings, both types of positive correlation are likely to be present, and sometimes the bias cancels out. For example, as we discuss in Section

<sup>13</sup>Another approach is to derive bounds on the deadweight loss from Equation 7. We have constructed analytical bounds based on properties of quadratic forms and their eigenvalues, but they will be informative only in special cases. This may be a promising area of future research.

5, cars that are strong substitutes are relatively likely to have similar fuel economy and similar durability.

The results above are directional and qualitative. In many cases, as we illustrate in the applications, simulation of the true welfare gain using a range of plausible demand elasticities can be highly informative. This usually does require some knowledge on the structure of the demand matrix, e.g. from existing empirical work in the literature.

## 2.5 Measurement error biases in $R^2$

To implement our method, the analyst needs estimates of the externalities (or other wedges). A practical concern in many settings will be accurate measurement of  $\phi_j$ . When inaccuracies in measures of  $\phi_j$  have a classical error structure it is straightforward to characterize the way that this biases the  $R^2$  statistic. An example of classical measurement error would be when  $\phi_j$  is estimated from a sample of observations, as in our first application, and sampling variability produces unbiased but variable estimates. Likewise, if  $\phi_j$  is an unbiased prediction, it may have classical mismeasurement. We use the following notation:  $\hat{\phi}_j \equiv \phi_j + \nu_j$ , where  $\phi_j$  is the true wedge,  $\hat{\phi}_j$  is the observed or estimated wedge, and  $\nu_j$  is therefore the error in measurement. Consistent with classical errors, suppose that  $\nu_j$  is independent of  $\phi_j$  (and of any regressors that are used in determining the tax scheme) and is distributed normally with mean zero and variance  $\sigma_\nu^2$ .

Suppose that the analyst regresses  $\hat{\phi}_j$  on  $f_j$ . This is a situation of errors in the dependent variable, so errors do not cause bias in the coefficients and the second-best policy is still consistently estimated. The  $R^2$  statistic, however, is biased downward due to the noise from mismeasurement. A simple derivation (see [Majeske, Lynch-Caris, and Brelin-Fornari \(2010\)](#)) shows that, in expectation:

$$\hat{R}^2 = R^2 \left( 1 - \frac{\sigma_\nu^2}{\sigma_\phi^2} \right) \quad (14)$$

where  $\hat{R}^2$  is the result from the estimation using mismeasured data and  $\sigma_\phi^2$  is the variance in  $\hat{\phi}_j$ . In terms of welfare interpretations, it implies that the second-best constrained policy will have larger welfare gains than indicated by the estimated statistic. In practical terms, where measurement error is a concern and errors are classical, an analyst can inflate the  $R^2$  upwards given an estimate of the signal-to-noise ratio in the data.

Another relevant source of noise in our setting could arise from the use of micro-data in performing the regression. Conceptually, the regression we have in mind should be run at the same level of detail over which the policy is being applied; if the benchmark policy differs with product  $j$  then a dataset containing different individual observations for each product should be collapsed to the product level before computing  $R^2$ . However, it is still possible to adjust the  $R^2$  from a regression run on the micro-data to recover the relevant welfare statistic. The average variance in the micro-data across products,  $\frac{\sum \sigma_j^2}{J}$ , can be substituted into Equation 14 in place of  $\sigma_\nu^2$ . The resulting

adjustment produces a value that is computationally equivalent to the  $R^2$  from the product-level regression, reflecting the welfare statistic we have in mind.<sup>14</sup>

## 2.6 Summary: when is our theory applicable?

The central point of our theory is that simple regression statistics often contain intuitive information about the welfare properties of corrective policies that face some design constraint. Figure 2 provides a visual summary of the situations under which our results obtain, which is intended to serve as an initial guide for those considering our methods in other applications.

All of our theory assumes that in the absence of policy, consumption of a good deviates from the social optimum due to some wedge  $\phi$ , such as an externality. Our base assumptions are that these wedges are fixed with respect to prices and that demand and supply are linear over the relevant range, as is generally assumed in the analysis of Harberger triangles. Under those assumptions, Equation 8 expresses the deadweight loss of an arbitrary vector of taxes that deviates from the Pigouvian benchmark. When demand and supply are not locally linear, it is possible to amend our results through simulation, which we illustrate in our electricity-pricing application.

When the conditions of Assumption 3 are met, then our results about the interpretation of the sum of squared residuals and the  $R^2$  will hold (Proposition 1). We interpret our first two applications, to electricity pricing and noisy laboratory measures, as meeting these condition most closely.

Even when violations of the assumptions are significant, the  $R^2$  may be a useful bound. As described in Proposition 3, particular types of correlations between tax errors and demand will create predictable bias in the  $R^2$  as a measure of welfare gain. In our vehicle longevity application, we demonstrate the size of this bias after introducing correlations calibrated from the literature.

In some cases, the cross effects will not simplify, and the bias will not fit the special cases embodied in Proposition 3. In that case, we suggest two approaches. One is to look for a modified relationship between regression statistics and welfare. This is what we do in application 4. There, we argue that  $R^2$  will be substantially biased, but that an alternative set of assumptions appropriate to the setting imply that the within- $R^2$  from a fixed effects regression has the desired interpretation (Proposition 4). Other approaches that incorporate additional market failures, endogenize  $\phi$  or

---

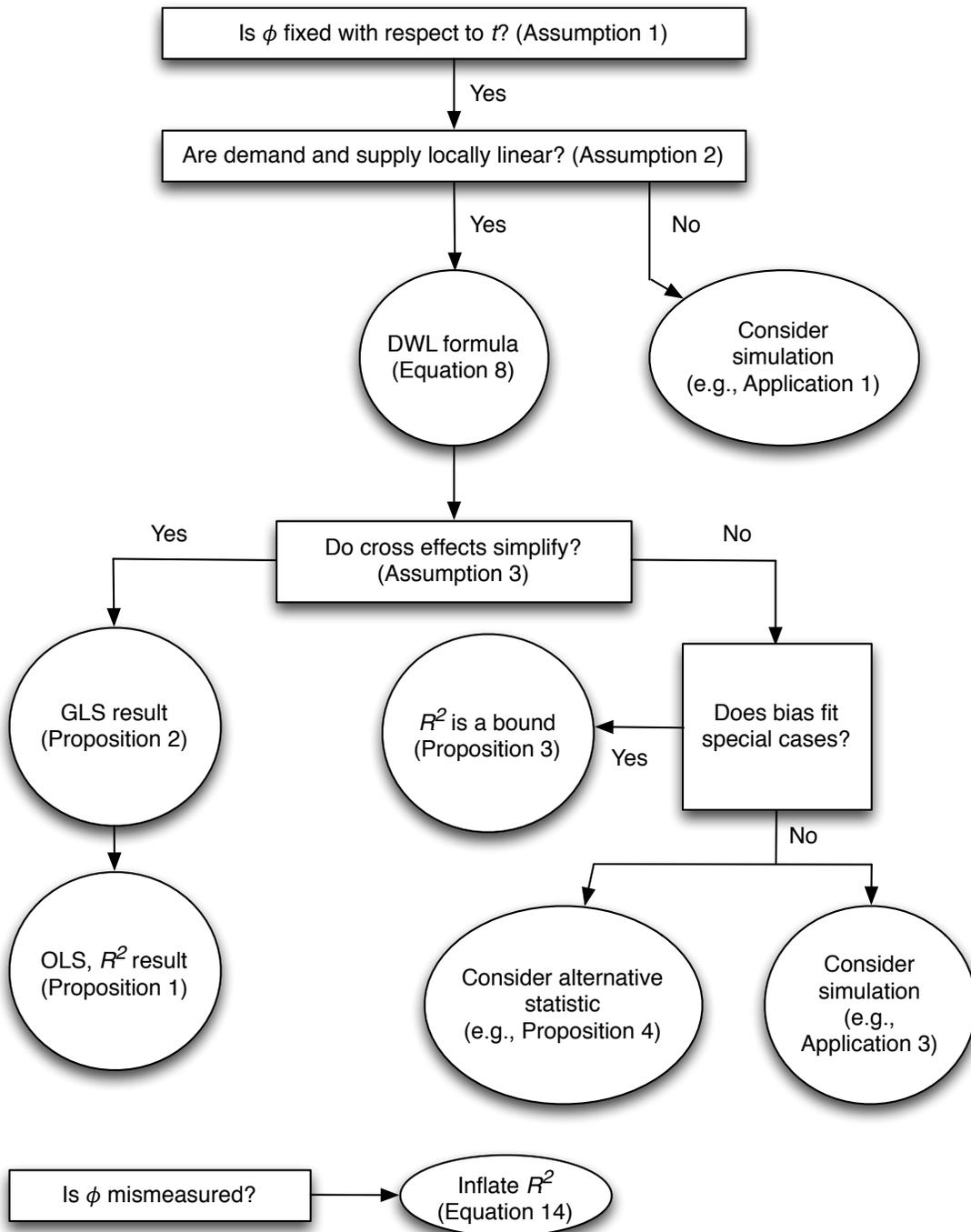
<sup>14</sup>Using micro-data adds noise to the regression and can be corrected using Equation 14. We show here why the relevant correction factor substitutes  $\frac{\sum \sigma_j^2}{J}$  for  $\sigma_\nu^2$ , where we define  $\sigma_j^2$  as the variance across observations of product  $j$ :  $\sigma_j^2 = \frac{\sum_{i=1}^I \phi_{ij}^2}{I} - \phi_j^2$ .

First note that in the general setting we define  $\sigma_\phi^2 = \sigma_\phi^2 + \sigma_\nu^2$ , where  $\sigma_\phi^2$  is the variance of the true product-level externality  $\phi_j$ . The variance of a micro-data set,  $\phi_{ij}$ , containing  $I$  independent observations on each of the  $J$  products can be written:

$$\sigma_{\phi_i}^2 = \frac{\sum_{j=1}^J \sum_{i=1}^I \phi_{ij}^2}{IJ} - \overline{\phi_{ij}^2} = \frac{\sum_{j=1}^J (I\sigma_j^2 + I\phi_j^2)}{IJ} - \overline{\phi_{ij}^2} = \frac{\sum_{j=1}^J \sigma_j^2}{J} + \frac{\sum_{j=1}^J \phi_j^2}{J} - \overline{\phi^2} = \frac{\sum_{j=1}^J \sigma_j^2}{J} + \sigma_\phi^2$$

The corresponding correction using (14) then raises  $R^2$  to the value it would take in a regression where the averaged data,  $\phi_j$ , is used on the left hand side. To the extent there is still important mismeasurement in the averaged data (for example sampling error in the form  $\frac{\sigma_j^2}{I}$ ), an additional correction can be layered on to the result here.

**Figure 2:** Schematic of Theoretical Results



consider other relaxations of our assumptions are key topics for future research. The other approach is to use simulation to determine whether calibrated degrees of correlation between tax errors and the demand system indicate that the bias in the  $R^2$  will be small or large. We demonstrate this approach in our vehicle longevity application.

Finally, at the bottom of Figure 2 we call out a practical consideration. Where there is classical mismeasurement of the wedges,  $R^2$  will be biased downwards. This can be corrected where information is available on the degree of noise in the data (Equation 14).

### 3 Application 1: Noisy Energy Efficiency Ratings

One 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.<sup>15</sup> This creates inefficiencies, and our theoretical 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 does 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 Corporate Average Fuel Economy (CAFE) program, which mandates that each firm meet a minimum average sales-weighted fuel economy of vehicles. 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. The test procedure established in 1978 included two courses; one each to represent urban and highway driving. 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. 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.

---

<sup>15</sup>Such mismeasurement naturally also occurs for non-energy goods and externalities. For example, to help prevent obesity, calorie labeling on menus will be mandatory for many U.S. restaurants. [Urban, McCrory, Dallal, Krupa Das, Saltzman, Weber, and Roberts \(2011\)](#) found using lab tests that while menus are, on average, pretty accurate, substantial variation exists. About 20 percent of foods purchased had at least 100 more calories than what was reported.

Yet again, the EPA instituted a new test procedure in 2008 that changed the ratings substantially on average, and also more for some vehicles than others.<sup>16</sup>

For political reasons, however, the CAFE program continues to use the less accurate original rating system from 1978.<sup>17</sup> 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 theoretical framework to quantify the welfare costs of using the old rating system in lieu of the updated one, via simple linear regression. 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. In other words, a sophisticated regulator can take out overall bias/tilt in measurement, but does not observe car-specific mistakes. These assumptions likely hold in practice—reported on-road fuel economy is close to the 2008 EPA ratings and the EPA explicitly presents differences between window sticker and regulatory CAFE fuel-economy values.<sup>18</sup>

### 3.1 The Pigouvian benchmark versus the constrained policy

In this application, we take a simplified view of the externalities associated with fuel economy and assume that the externalities associated with an automobile are proportional to its true fuel consumption per mile. This is consistent with how fuel-economy standards have been designed, as such standards impose a shadow cost on each vehicle equal to a linear function of the vehicle’s fuel economy rating (see, e.g., [Anderson and Salée 2016](#)). (In application 3, we challenge this notion and discuss various complications, but here we wish to focus only on the issue of mismeasured test ratings, not other problems with fuel-economy regulation.)

In terms of our model, each product  $j$  is a type of car. The externality  $\phi_j$  is some factor  $\zeta$  (e.g., the social cost of carbon times carbon emissions per gallon of gasoline times total miles driven) times true fuel economy. The shadow taxes imposed by CAFE will be a linear transformation of the old ratings, which the regulator is constrained to use in setting policy:

$$\begin{aligned}\phi_j &= \zeta \times \text{New Fuel Economy Test Rating}_j \\ t_j &= \alpha + \beta \times \text{Old Fuel Economy Test Rating}_j.\end{aligned}$$

---

<sup>16</sup>This procedure involved five separate dynamometer tests – the original two tests and three new ones. Several tests are combined to determine the highway and city ratings that appear on fuel-economy labels for consumers.

<sup>17</sup>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.

<sup>18</sup>See <http://www.epa.gov/fueleconomy/documents/420f14015.pdf> and <http://www.epa.gov/fueleconomy/documents/420b14015.pdf>.

### 3.2 Will cross effects simplify?

Where the noise in measurement is uncorrelated with factors that determine vehicle demand, the weaker version of Assumption 3 from Section 2 will hold, and the main theoretical results in Proposition 1 and Corollary 1 apply. It is logical to suppose that errors in the tax rates from an unbiased policy are uncorrelated with cross-price derivatives, as they are likely to be due to idiosyncratic aberrations from test trials or particular technologies, like stop-start systems, that are of little concern to consumers (and therefore not correlated with cross-price derivatives).

In this case, the  $R^2$  from a regression of  $\phi_j$  on  $t_j$  has a welfare interpretation. 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 (the second-best) in place of the accurate ratings (the Pigouvian benchmark). Note, however, that because  $\phi_j$  is proportional to the old fuel-economy ratings and  $t_j$  is a linear transformation of the new fuel-economy estimates, that the  $R^2$  of interest is identical to the  $R^2$  from a regression of the new fuel-economy rating on the old one.

### 3.3 Data

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.<sup>19</sup>

### 3.4 Results

We plot these data in Figure 3. 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.<sup>20</sup> 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.<sup>21</sup> This indicates that, along the dimension of test rating quality, the efficiency gain from removing noise is quite minor. The vast majority of the welfare gain from an optimally designed fuel-economy policy that used the new ratings can be achieved by a policy that uses the old rating system. Interestingly, this

---

<sup>19</sup>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.

<sup>20</sup>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.

<sup>21</sup>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 3:** Old and New Combined Fuel-Economy Ratings

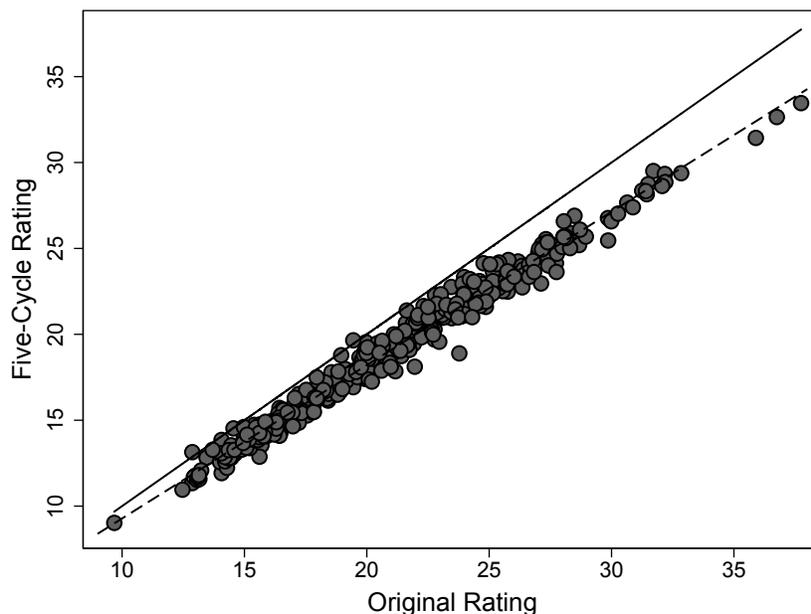


Figure shows the pre-2008 (original) combined fuel-economy rating and the post-2008 (five-cycle) rating in miles per gallon for a sample of vehicles. Dashed line is linear fit. Solid line is the 45-degree ray.

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 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 durability discussed in Section 5.

## 4 Application 2: Real-Time Electricity Pricing

Our second application is to time-varying electricity prices. Electricity consumers typically pay the same price for electricity regardless of when they consume it. In contrast, the marginal cost of producing electricity varies significantly across hours of the day, days of the week, and months of the year due to variance in the marginal source of generation. At low levels of demand, marginal cost is low because only solar, wind and so-called inexpensive “base load” power plants are needed. At high levels of demand, higher cost “peaker” plants produce the marginal unit. As a result, the marginal cost of electricity is frequently several times higher at one hour of the day as compared to another hour in the same day.

Economists have contemplated the efficiency benefits of time-varying pricing schemes that align

price and marginal cost. The theoretical ideal is called real-time pricing, which is a scheme in which the price of electricity charged to the consumer is unconstrained and is adjusted at a high frequency to reflect costs. Real-time pricing provides the right incentive to consumers at every moment and therefore achieves the efficient resource allocation, provided that no other markets failures are present (Borenstein and Holland 2005).<sup>22</sup>

Historically, it was infeasible to measure electricity consumption hour by hour for each end user, so this mispricing was a necessary compromise. However, with the advent and roll out of computerized electricity meters, high frequency measurement at the customer level is already a reality in most parts of the United States. Even so, real time electricity prices have met with considerable resistance from utilities and regulators, who fear that consumers will complain about price surges and unpredictable bills.

As a result, while the technology to implement real-time pricing is already in place, pricing reforms have been incremental. Instead of real-time pricing, utilities have experimented with peak pricing for certain times of the day, seasonal rates, or peak prices only on certain days on which demand is forecasted to be very high due to weather. A significant literature in economics has evaluated these programs, primarily with a focus on how demand responds to price variation (Jessee and Rapson 2014; Andersen, Hansen, Jensen, and Wolak 2017; Fowle, Wolfram, Spurlock, Todd, Baylis, and Cappers 2017; Gillan 2017; Ito, Ida, and Tanaka 2018). A remaining unanswered question in this literature is whether most of the efficiency gains from real-time pricing can be achieved by these intermediate policies. If simpler rate designs can capture most of the efficiency gains of real-time pricing, then this may present a useful way forward for the industry that can accelerate reform.

We demonstrate that our model can be used to answer this policy-relevant question with readily available data and simple OLS regressions. We use wholesale pricing data from a major electricity market in the Eastern U.S., which provides a measure of the marginal cost of electricity at the hourly level. Using our model, we show that the  $R^2$  from a regression of observed wholesale prices on season, day-of-week, or peak demand periods measures the proportion of the welfare gain that an intermediate reform that allows tariffs to vary by those variables would achieve, relative to the welfare gain that would be achieved by moving all the way from a flat rate to real-time pricing. As such, our method allows us to evaluate a wide range of alternative policies with minimal effort. This can be quite valuable because the welfare gain achieved by intermediate policies will vary based on the characteristics of supply and demand, e.g., demand variability and capacity constraints in the system. This increases the value of being able to quickly calculate potential welfare gains across a number of different markets in terms of both time and geographic scope.

In our application below, we find that the intermediate schemes perform relatively poorly. Fairly complex schemes are required to recover half of the welfare gains from real-time pricing, and schemes that mimic real-world policies used to date recover only a small fraction of the potential

---

<sup>22</sup>Although not the focus of this application, pollution externalities can be introduced to our analysis by adding environmental damages to private marginal costs and running the regressions with social marginal costs as the dependent variable.

gains. These results should prove useful in the ongoing debate about electricity rate design, which is poised to undergo significant reform in the coming years.

The insights from this application will also apply to other settings that feature coarse pricing, where many related goods must be given a common price due to some exogenous constraint on the pricing policy, even though social costs differ due to production technologies, scarcity or externalities. Potential examples include markets for parking, traffic congestion, taxis/ride-sharing services, or event tickets.

#### 4.1 The real-time pricing benchmark versus the constrained policy

To apply our model to electricity, we interpret each product  $j$  as electricity consumed at a specific moment. Empirically, we will consider an hour to be a unique moment, because this is the granularity of our wholesale pricing data. We focus on a single integrated electricity market, so we do not need to consider electricity consumed at different locations to be different goods.<sup>23</sup>

In our model, consumers pay  $p_j + t_j$  where  $t_j$  is understood as a tax and  $p_j$  is a uniform producer price. Here, we interpret  $t_j$  as the tariff that applies to good  $j$ , so that the final price to consumers is just the tariff  $t_j$  (equivalent to assuming  $p_j = 0$  in the original notation). Under a flat tariff,  $t_j$  is the same across all  $j$  goods. Under real-time pricing, the tariff is unique to each  $j$ . Intermediary policies will have subsets of  $j$  (such as peak demand periods) for which consumers face a common tariff.

Unlike our other applications, there is no externality. Thus, the full social cost of producing a unit of good  $j$  is just the marginal cost  $mc_j$ , which we allow to vary across time. Any mismatch between the tariff and marginal cost induces an inefficiency, where the wedge is equal to  $t_j - mc_j$ . This wedge plays exactly the same role in our theory as the wedge due to imperfect correction of an externality (which is denoted  $t_j - \phi_j$  in the other applications). Thus, in terms of our model:

$$\begin{aligned}\phi_j &= mc_j \\ t_j &= \alpha + \beta' z_j.\end{aligned}$$

where  $z_j$  is a vector that includes tariff policy variables such as on versus off-peak or day-of-week indicators. Note that  $z_j$  can represent any tariff scheme that is linear in parameters, including interactions of indicator variables. The method can thus evaluate highly flexible tariffs.

If our assumptions about demand and supply hold, then the  $R^2$  of a regression of  $\phi_j$  on  $t_j$  will indicate the welfare fraction achieved by the constrained pricing scheme (second-best) relative to the real-time pricing benchmark, where both welfare gains are calculated relative to an unbiased

---

<sup>23</sup>As discussed further below, we use data from PJM, an integrated electricity market that spans multiple states. We treat this market as a single location and use PJM's reported system price. In reality, there are sometimes transmission constraints that imply that delivering electricity to one specific location has higher cost than delivering to another location at the same time, even within the market. Most U.S. electricity markets therefore have locational marginal prices (LMPs) that are specific to a particular node in the grid. We abstract from this issue, as does the bulk of the literature. In principle, however, our method could be used to gauge the welfare implications of the granularity of prices across geographical space as well as the time dimension focused on here.

flat tariff.<sup>24</sup> Even the real-time pricing benchmark is not quite first-best since it is granular on hour, ignores transmission constraints, etc. Thus, as usual, we measure the efficiency gain from more granular pricing along the dimension that policy makers can realistically target; in this example: average hourly tariffs.

## 4.2 Will cross effects simplify?

In this application increasing marginal costs are essential. As discussed in Section 2, our results still apply in this case, but the assumptions should be interpreted in terms of combined responses of demand and supply. As detailed in Appendix A, zero cross-price derivatives (the strong version (b') of Assumption 3) for demand and supply is sufficient for all our results from Section 2 to go through. In that case, we can characterize the deadweight loss of using an arbitrary vector of tariffs, denoted  $t_j = \tau_j$ , as compared to using real-time pricing, as the sum of  $J$  Harberger triangles:

$$-2 \times DWL(t = \tau) = \sum_{j=1}^J (\tau_j - mc_j)^2 \frac{\partial \tilde{x}_j}{\partial t_j},$$

where  $\frac{\partial \tilde{x}_j}{\partial t_j} = \frac{\partial x_j}{\partial t_j} - \frac{\partial mc_j}{\partial t_j}$ . Minimizing this distortion will involve fitting the tariff schedule so as to minimize the sum of squared errors between the tariff and the observed marginal cost, weighted by the derivative terms. When  $\frac{\partial \tilde{x}_j}{\partial t_j}$  is uncorrelated with the wedges or common across all  $j$ , then the formula simplifies to its final form and our  $R^2$  result applies.

Does it make sense to assume that cross-price derivatives are zero for supply and demand? On the demand side, the required assumption is that a change in the tariff in hour  $j$  does not affect demand in hour  $k \neq j$ . Substantial empirical support exists for this assumption. A consistent finding in the literature is that such cross-price derivatives are quite small, and are often statistically indistinguishable from zero. In other words, the electricity tariff during hour  $j$  does not affect the demand for electricity during hour  $k \neq j$ .

Specifically, a substantial literature has studied experiments that raise the cost of electricity at specific hours of the day; for example on weekdays during late afternoons in the summer, when system demand peaks due to air conditioner use in homes. A common question has been to what extent consumers will reduce electricity consumption during this high-price window and substitute this for consumption in “shoulder” hours around the experiment. Such studies consistently find that peak tariff schemes lower consumption during the targeted window but reveal minimal shifting of demand into off-peak hours (Jessee and Rapson 2014; Fowle et al. 2017; Gillan 2017; Ito, Ida, and Tanaka 2018). The exception is Andersen et al. (2017) which finds that a variable pricing scheme does cause significant shifts of demand into lower priced windows in Denmark. Therefore, the strong version (b') of Assumption 3 is likely appropriate in this application, at least in many

---

<sup>24</sup>It may seem counterintuitive that we can use historical data that come from observed marginal costs, even though those realized marginal costs depend on the particular flat tariff that was in place during the sample. Appendix A shows that, under local linearity, the  $R^2$  of a regression of observed marginal costs (under the flat tariff) on the policy variables equals the  $R^2$  of a regression of the benchmark marginal costs (under real-time pricing) on the policy variables. Hence, the relative efficiency gain can be computed from a regression that directly corresponds to our data.

circumstances. Bolstered by this evidence, we proceed by assuming that cross-derivatives are zero, but we also assess the performance of  $R^2$  using estimates from the literature that quantify how large cross effects might be in Section 4.5 below.

On the supply side, the corollary question is whether the price of electricity in time period  $j$  affects the cost of production in time period  $k \neq j$ . It is reasonable, and indeed common in the literature, to assume that production costs in different hours are separate production processes and are not directly related. Marginal costs are likely to be serially correlated, but this is because demand is serially correlated not because production in one time causes a shift in cost in other hours.<sup>25</sup> (Recall that we allow the marginal cost of production to be rising at any moment  $j$ . The assumption is that price in one hour does not affect cost in a different hour.)

### 4.3 Other modeling considerations

Throughout our theory, we maintain the assumption that demand and supply curves are linear over the relevant range of prices. Note that we only require that each good  $j$  has locally linear supply and demand, not that the demand and supply curves across products  $j$  have the same slope. Local linearity does not seem like a problematic assumption on the demand side, but electricity supply curves can become convex when a market approaches capacity limits (though we show empirically that a linear supply assumption still fits a large part of the supply curve). When supply is convex, the  $R^2$  statistic will provide an approximation. We investigate its accuracy via simulation, given data on the shape (convexity) of the market-level supply curve. In our case, the approximation appears to be quite good (see Section 4.5 below).

### 4.4 Data

Our empirical application uses data from the PJM wholesale electricity market. While originally comprising the states of Pennsylvania, New Jersey and Maryland (thus, the name PJM), the PJM market is a regional transmission organization (RTO) that runs one of the largest wholesale electricity markets in the U.S., stretching into 13 states in eastern and central U.S. plus the District of Columbia. PJM is one of five RTOs that run an active market for wholesale electricity. As in most wholesale electricity markets, PJM runs an hourly real-time auction for energy, bringing together producers and consumers (typically, utility companies) of electricity. These auctions yield hourly wholesale prices for electricity, which are a good measure of marginal costs. We use hourly pricing data for the year 2012, as this is the year for which we have data for the supply curve which we use in the convexity simulation below.

---

<sup>25</sup>Technically this may not be true for adjacent hours because of startup and ramping costs for fossil-fueled plants. We follow much of the literature in assuming that their impact on key results is modest, though we note that [Reguant \(2014\)](#) and [Cullen \(2015\)](#) are exceptions that model startup costs explicitly.

## 4.5 Results

In this section, we report the  $R^2$  for a wide variety of alternative pricing schemes. We also discuss the results from simulations that introduce cross-derivatives and convex supply (details are in Appendix B). We run variants of the following regression:

$$price_{th} = \alpha + \beta' z_{th} + \varepsilon_{th},$$

where  $t$  indexes date,  $h$  indexes hour,  $price_{th}$  is the observed wholesale electricity price,  $z_{th}$  is a vector that includes potential tariff policy variables such as on vs. off-peak indicators  $\theta_p$ , hour of day fixed effects  $\theta_h$ , day of week fixed effects  $\theta_d$ , monthly fixed effects  $\theta_m$ , season fixed effects  $\theta_s$ , or their interactions.

Table 1 shows the efficiency gain from using increasingly flexible tariff policies. We find that simple, yet commonly used, tariff structures like on vs. off-peak prices do not improve efficiency much. Even the highly sophisticated—and potentially hard to understand for consumers—pricing schemes that we analyze (such as tariffs that vary by weekday, hour and month) capture less than half of the efficiency gain of real-time electricity pricing. This conclusion is similar to Borenstein (2005), who uses a more detailed simulation model of competitive electricity generation.

**Table 1:**  $R^2$  from Electricity Tariff Regressions

Pricing Regime	$R^2$
On vs. off peak fixed effects	0.040
Hour of day fixed effects	0.135
Hour of day & day of week fixed effects	0.153
Hour of day & month of year fixed effects	0.193
Hour of day, day of week, & month of year fixed effects	0.211
Day of week interacted with month of year fixed effects	0.297
Hour of day interacted with day of week interacted with month of year fixed effects	0.428
Number of observations	8,784

*Notes:* Dependent variable is the hourly price of electricity observed in the PJM market for 2012. Peak hours are defined as 2-6 p.m.

We test the robustness of this finding by assessing the performance of  $R^2$  when cross-price derivatives in demand are not zero. Andersen et al. (2017) report the largest cross-price derivatives among the studies we discuss above, with substitution to shoulder hours of approximately 29% of the own-price effect. Appendix B.1 compares the  $R^2$  with welfare calculations that account for such substitution. Even with considerable substitution, the largest bias in the  $R^2$  measure is approximately two percentage points. We also consider the cross-price effects estimated in Ata, Duran, and Islegen (2016), as well as cases with spillovers across hours as in Jessoe and Rapson (2014). We find similarly small biases.

Finally, we evaluate how our results change when we take into account that electricity supply is convex at high levels of capacity utilization. To do this, we make use of plant-level engineering data from the same year. These data report both capacity and engineering marginal costs for each

plant in the system, forming a step-function supply curve.<sup>26</sup> We capture the convexity of supply in two ways. First, we estimate a quadratic supply curve through the aggregate engineering marginal cost curve. Second, we use the actual step function.

We then simulate the welfare gains from each of the seven pricing regimes above and compare these to their respective  $R^2$  measures. We find that  $R^2$  remains a reliable indicator for the efficiency gain of constrained policies even when marginal costs are convex.  $R^2$  is always within 1% of the simulated welfare gains under our quadratic estimate of supply and generally within 10% using the step-function supply curve. Furthermore, the bias in the  $R^2$  when using the step-function supply curve is relatively constant across pricing regimes suggesting that comparisons across different pricing regimes can still be made. See Appendix B.2 for details.

## 5 Application 3: Automobiles and Longevity

The total externality caused by an energy-consuming durable good depends on both its energy efficiency and its lifetime utilization; e.g., a car’s lifetime gasoline consumption depends on fuel economy and miles driven. Were all products utilized the same amount, a set of product taxes based only on energy efficiency could accurately target lifetime externalities, thereby shifting demand across products efficiently. But, heterogeneity in the longevity of products with the same energy efficiency rating implies that energy efficiency policy is necessarily imperfect.

We demonstrate the empirical importance of this issue using the case of greenhouse gas emissions from automobiles. We use a novel data set to estimate the average lifetime mileage of different car models, and we translate that into lifetime damages from greenhouse gas emissions, according to each vehicle’s fuel economy and the social cost of carbon. We then use simple linear regressions motivated by our theory to evaluate second-best policies that must construct a tax vector for vehicles that depends on a vehicle’s fuel economy, but not its longevity. As we explain below, this constrained policy closely resembles the dominant real-world policy in this sector, fleet-average fuel-economy standards, such as the U.S. CAFE program.

We find that the constrained policies are highly inefficient. There is a voluminous literature that explores the welfare implications of energy-efficiency policies, but we know of no prior paper that has demonstrated the importance of heterogeneity in product longevity.<sup>27</sup> We speculate that this issue is a first-order concern not just for automobiles, but also appliances, building codes and other efficiency programs.

### 5.1 The Pigouvian benchmark versus the constrained policy

We study the efficiency of a fuel-economy standard in correcting greenhouse gas emissions in the purchase of new vehicles. In terms of our model, each type of vehicle is a product  $j$ . We model the

---

<sup>26</sup>Data such as these have been used extensively to calculate hourly equilibria within electricity markets.

<sup>27</sup>Allcott and Greenstone (2012) note that differences in utilization might justify geographically differentiated appliance standards, but they do not quantify heterogeneity or calculate potential gains from differentiation.

lifetime greenhouse gas related externality from automobiles as proportional to the total gasoline consumed by each vehicle type  $j$ . Total gasoline consumed is the total lifetime mileage of a vehicle multiplied by its miles per gallon efficiency. To translate this into a dollar externality, lifetime emissions are multiplied by a constant,  $\psi$ , which is equal to the social cost of carbon per gallon of gasoline. Note that a gasoline tax would achieve the Pigouvian benchmark, so long as consumers are aware of product durability and have a rational forward-looking valuation of fuel costs.

In contrast, a fleet-average fuel-economy standard will create a shadow tax scheme where shadow taxes are equal to a linear function of a vehicle’s fuel-economy rating.<sup>28</sup> Thus, in terms of our model:

$$\begin{aligned}\phi_j &= \psi \times \text{Fuel Economy Rating}_j \times \text{Lifetime Mileage}_j \\ t_j &= \alpha + \beta \times \text{Fuel Economy Rating}_j.\end{aligned}$$

If our assumptions all hold, then the  $R^2$  of a regression of  $\phi_j$  on  $t_j$  will indicate the fraction of the welfare gain achieved by a fuel-economy standard (second-best tax), which depends on only fuel-economy ratings, as compared to the welfare gain achieved by a policy based on both fuel economy and lifetime usage (the Pigouvian benchmark), where both gains are measured compared to a policy that places a uniform tax on all cars equal to the average externality. In Section 5.3, we discuss additional complications related to vehicle-related externalities and ways in which fuel-economy regulation differs from gasoline taxation that are not captured in this description.

Note that our unit of observation is the vehicle model (e.g., a 1995 Toyota Corolla). Due to accidents and random mechanical failure, individual units will be scrapped with variable lifetime mileage, but this is orthogonal to our welfare comparisons. That is, all of our results are robust to allowing for random product failure, with  $\phi_j$  interpreted as the mean externality—so long as the random failure rates are not endogenous to product taxes (which we return to below). The reason is that ex ante unknowable variation in damages across identical units cannot be targeted by any new vehicle policy, so this will not affect our comparison of second-best policies to the Pigouvian benchmark.

## 5.2 Will cross effects simplify?

Our main assumptions are about cross-price derivatives across types of automobiles. In this market, cross-price derivatives are clearly important, so the strong version of our Assumption 3(b’) will not hold. Instead, we argue that cross-effects will plausibly be small and will cancel out in the  $R^2$  ratio under a second-best tax policy so long as vehicles that have greater or lesser longevity, conditional on fuel-economy ratings, are not systematically closer or further substitutes for each other. In that case, Assumption 3(b) will hold. But, we relax that assumption empirically by allowing vehicles

---

<sup>28</sup>Historically, policies like CAFE were firm-specific, so that the shadow price varied across firms. CAFE, and most similar policies in other countries, now allow trading, which fits our description here. However, there is suggestion that trading in CAFE has been thin (Leard and McConnell Forthcoming), but this may be because trading was introduced alongside footprint-based standards, which reduces the variation in shadow costs across firms (Ito and Sallee Forthcoming).

with more similar longevity and more similar fuel economy to be closer substitutes, following the derivations in Section 2.4. Introducing these correlations turns out to have limited impact; our results are robust.

### 5.3 Other modeling considerations

Below we find that second-best constrained policies are highly inefficient. We interpret this as evidence that fuel-economy regulations are inefficient as compared to a gasoline tax. But, this interpretation is generous to fuel-economy regulation, because it abstracts from other well-known inefficiencies in fuel-economy policies. In particular, fuel-economy regulations fail to incentivize abatement on the intensive margin; e.g., a fuel-economy standard can get people to buy the optimal vehicle, but they will not drive the optimal number of miles. Our model abstracts from that by assuming that the externality attached to each vehicle is fixed. Note, however, that we are concerned with lifetime mileage, so the intensity-of-use margin that concerns us is only the scrappage decision, not miles traveled per year.<sup>29</sup> In addition, revenue-neutral energy-efficiency policies fail to get the average price of goods right; e.g., a fuel-economy standard can get the relative price of inefficient versus efficient cars right, but all cars will be too inexpensive and the car market will be too large overall.<sup>30</sup>

Our welfare analysis considers two alternative tax structures, a second-best scheme that imposes a tax on each vehicle that is a linear function of its fuel economy rating, and a Pigouvian benchmark that imposes taxes according to each vehicle’s externality. As such, we measure welfare loss along the vehicle purchase margin, which is directly targeted by vehicle-based taxes. This abstracts from the market size effects (by assuming both tax schedules are correct on average) and the intensive margin effect (which is omitted from both policies and therefore likely has limited impact on the proportional gains we emphasize), and bases the policy comparison only on differences related to tax rate errors driven by heterogeneity in longevity.

Thus, our  $R^2$  results can be interpreted as an upper bound on the fraction of the welfare gain from a gasoline tax that can be achieved by a second-best fuel-economy regulation. It is an upper bound because a gasoline tax would also achieve gains along the scrappage (intensity of use) margin, and because a gasoline tax would correct the overall size of the car market by raising the average price of automobiles.

The Pigouvian vehicle tax does neither. In brief, our comparison—within which CAFE performs quite poorly—understates the real welfare losses incurred from using CAFE instead of a tax on gasoline.

Finally, recall that our model is for a representative consumer. Individual drivers may have different on-road fuel consumption rates for identical cars due to differences in driving styles and conditions (Langer and McRae 2014). Differences in maintenance or accident risk may imply

---

<sup>29</sup>See Jacobsen and van Benthem (2015) for evidence on how scrappage decisions influence the welfare implications of fuel-economy regulations.

<sup>30</sup>See Holland, Hughes, and Knittel (2009) for an exploration of how performance standards create inefficiencies due to their average price effects.

that some drivers “use up” a vehicle faster than others, so that expected lifetime mileage for a car depends on driver behavior. We abstract from these considerations, both because of data limitations and because we doubt their quantitative significance. As discussed above, we are not concerned with random failure that is unpredictable to the consumers themselves at the time of purchase. Thus, driver heterogeneity is relevant only to the extent that different types of drivers sort into different vehicles systematically in response to changing taxes. Moreover, our model does permit heterogeneity in miles driven per year—all calculations are done in terms of total miles driven from new until scrappage, regardless of calendar age. Heterogeneity in annual usage matters only to the extent that faster or slower rates of utilization affect the total expected lifetime mileage of the vehicle. The fact that most cars have several owners over their life will tend to de-couple any individual owner from the vehicle and will mitigate concerns related to individual heterogeneity.

## 5.4 Data

Our data come from the California Smog Check program, which records the odometer reading for all tested vehicles. We merge these data with a national registration database that identifies when a vehicle has been retired from the U.S. fleet, and take the last observed odometer reading before a vehicle’s retirement as the measure of its lifetime mileage. We aggregate individual observations to the VIN10 level (the finest distinction of a unique car type possible in our data, which delineates a vehicle by make, model, model year, engine size and, often, transmission, drive type, body style and trim) and VIN8 level (which encompasses the same vehicle characteristics as the VIN10 but aggregates across model years). We divide lifetime mileage by official fuel-economy ratings to estimate lifetime gallons consumed.<sup>31</sup> We use this as our measure of the lifetime externality of each vehicle type (i.e.,  $\phi_j$ ), multiplying by the social cost of carbon per gallon of gasoline when necessary to convert the externality into dollars.<sup>32</sup>

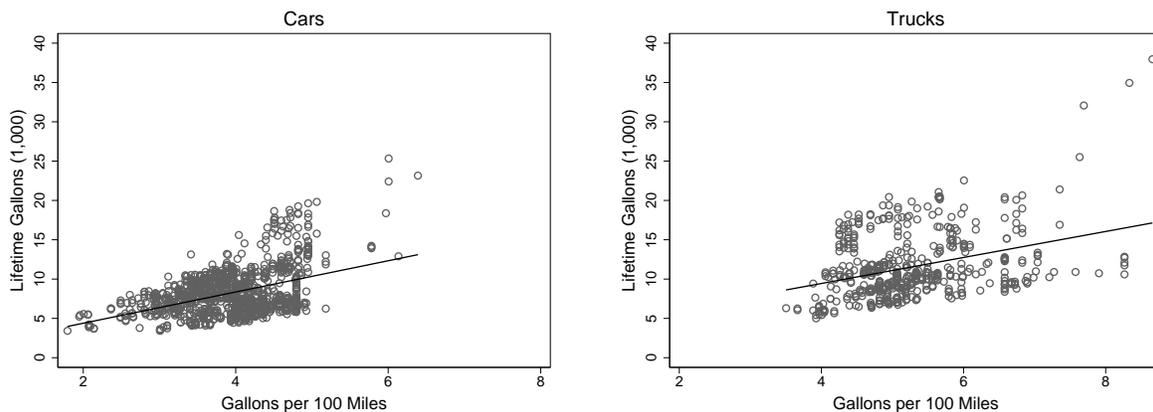
We do not observe all units, which creates the possibility of measurement error and censorship bias. Regarding the former, we are concerned with the average mileage at scrappage of cars but we observe only a sample. To the extent this error may be large the  $R^2$  value can be adjusted following the discussion in Section 2.5. However, we demonstrate that for our sample the bias is likely to be very small (see Section 5.5 and Appendix C). Regarding censorship bias, we do not observe cars under six years old (as they are usually not required to be tested), cars that were retired before our data began, or cars that were still in the fleet when our data ended. These missing observations could bias our results in either direction, but we demonstrate through several procedures that this censorship bias is apparently small and likely causes us to slightly overstate

---

<sup>31</sup>This 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. This means that social cost growth offsets discounting.

<sup>32</sup>We abstract from carbon emissions related to construction and scrappage of vehicles because standard estimates suggest that these emissions make up only 8% of life cycle emissions (National Research Council 2010). The remainder is due to gasoline consumption. Moreover, to the extent that these emissions are the same across models, incorporating them would have no effect on our welfare calculations. Only heterogeneous life cycle emissions matter.

**Figure 4:** 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 and to model years 1988 to 1992. Observations with VMT above one million miles are dropped. Solid lines are OLS prediction lines.

the efficiency of fuel-economy policies. We use comprehensive national registration data from R.L. Polk to address concerns related to missing data. More data details and robustness checks are included in Appendix C.

## 5.5 Results

In this section, we report the  $R^2$  from several alternative specifications. In Appendix C we explore further robustness checks. In Section 5.6, 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.

We begin by showing the data for our preferred sample in a scatterplot. Figure 4 shows a scatterplot of the relationship between a vehicle's total lifetime externality (gallons of gasoline) and its official fuel consumption rating, for both cars and trucks, along with the OLS fitted line. A point in the figure corresponds to the average lifetime 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 (more on this below), and to vehicle models for which we have at least 200 observed retirements. We drop observations with more than one million miles to limit the influence of outliers.

There is, as expected, a 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 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

**Table 2:** Regression  $R^2$ 

	VIN10-prefix		VIN8-prefix	
	OLS	WLS	OLS	WLS
<b>All model years</b>				
All models	.26	.20	.23	.19
Models with $N \geq 200$	.22	.17	.27	.19
<b>Model years 1988-1992</b>				
All models	.27	.26	.28	.27
Models with $N \geq 200$	.29	.22	.34	.25

Note: Table shows  $R^2$  from regressions using VIN-prefix average lifetime gallons consumed on fuel consumption rating. The unit of observation is either a VIN10-prefix or a VIN8-prefix in the first panel. Observations with  $VMT$  above one million miles are dropped.  $N$  is the number of observed retirements, and WLS weights the regressions by  $N$ .

sample regression is 0.29.) According to our theory, this implies that the second-best linear policy captures only 18% and 12% of the welfare gains for cars and trucks that would be achievable with an efficient set of product-based taxes that varies not only with fuel economy, but also with vehicle durability.

Note that the second-best policy will under-tax long-lived vehicles (observations above the regression line), and it will over-tax short-lived vehicles (observations below the regression line). Some may find it counterintuitive that the efficient policy would raise taxes on long-lived vehicles. To see the intuition, consider two vehicles with the same fuel-economy rating, where one lasts twice as long as the other. To drive the same number of miles (same emissions), two short-lived vehicles will be required, so that the tax will be paid twice. Thus, harmony between the tax paid and the emissions emitted requires taxing the long-lived vehicle more heavily.

Table 2 reports the  $R^2$  from a set of regressions that take the form:

$$\text{Average Lifetime Gasoline Consumption}_j = \alpha + \beta \text{Gallons per Mile}_j + \varepsilon_j, \quad (15)$$

where  $j$  indexes a vehicle type (VIN10-prefix or VIN8-prefix). We report a range of estimates in order to assess the importance of sample restrictions, weighting, censoring, the level of aggregation, and sampling error. Weighted least squares (WLS) results weight VIN-prefixes by the number of observed retirements  $N$ . These results are useful for assessing the effect of sampling variation, but they also approximate weighting by sales share, which leads to the preferred welfare interpretation because it aggregates to total externalities generated. 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 unit of observation is average gasoline consumption across vehicles with the same VIN10-prefix or VIN8-prefix, consistent with Figure 4 above.

Table 2 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 are 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, which are the years in our data with the least censorship concerns and therefore our preferred specification. 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) will cause the estimated  $R^2$  to be below the true welfare gain ratio. To assess the importance of sampling error, we compare results from OLS to WLS, which weights models by the number of vehicles 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 in Appendix C.

Our theoretical results are focused on second-best policies—tax schedules that are set optimally against some design constraint—but actual policies may deviate from the second-best. In Appendix C we comment further on biased policies, which can come either because the average tax is wrong (“mean bias”) or because the slope is wrong (“slope bias”).

### Summary of additional estimates

Our approach also applies to more flexible fuel-economy policies; the  $R^2$  from the appropriate regression will have the same welfare interpretation for any policy that is linear in parameters. For example, fuel-economy policy could put a shadow price on each model that was a quadratic function of fuel consumption ratings. Or, tax rates could be based on fuel-economy bins. Then, the  $R^2$  from a regression of the externality on fuel consumption and fuel consumption squared, or of discrete bin dummies, would have the desired interpretation. More flexible fuel-economy policy could also base shadow prices on not just fuel economy, but also other attributes, like class (car versus truck), model year or body style. The  $R^2$  from regressions of lifetime externalities on fuel consumption and these additional attributes directly indicates the welfare gains possible from more flexible policies.

We ran a number of such regressions and summarize the results here. First, we allow for separate fleetwide average standards for cars and light-duty trucks (similar in spirit to the initial structure of U.S. CAFE standards). In our framework, this corresponds to adding a truck indicator and its interaction with the fuel consumption rating. This adds very little explanatory power. In our preferred specifications using model years 1998-1992,  $R^2$  rises by 0.004-0.013 to a range of 0.23-0.35, depending on the specification in Table 2). This strongly rejects the notion that separate regulation of cars and trucks was useful in addressing the inefficiency that we identify. Adding other policy attributes such as body style and model year has a similarly small impact on the  $R^2$ ; none of these attributes are strongly correlated with durability (conditional on fuel economy).

We also considered attribute-based policies based on fuel consumption and vehicle size, either using the vehicle’s footprint (wheelbase  $\times$  track width, where wheelbase is the distance between the front and rear axles of a vehicle) or by more flexibly including wheelbase and width as separate regressors. In both specifications, we also allowed the policy to be different for cars vs. light-duty trucks. This mimics current U.S. CAFE policy, which is based on fuel consumption and footprint, with separate standards for cars vs. trucks. However, including footprint in our regressions has little effect. It raises  $R^2$  by 0.004-0.049 to 0.25-0.35. Including wheelbase and width has a somewhat larger effect:  $R^2$  increases by 0.077-0.096 to 0.31-0.43. This suggests that there could be efficiency gains from more flexible sized-based standards due to the correlation between size and longevity, though such standards create distortionary incentives (Ito and Sallee Forthcoming).

In all cases, the qualitative conclusion remains that there is substantial variation in lifetime consumption that is not explained by fuel economy, vehicle type, or size, which implies that policies based only on such vehicle attributes, 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).

## 5.6 Estimates of deadweight loss

We can translate the relative gains from the Pigouvian 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 1990 model year (typical of the years in Table 2 above), computing the possible welfare gains from a Pigouvian product-level tax and the deadweight loss from the second-best tax based on fuel economy. We then explore the influence of a range of substitution patterns across vehicles, following the theory in Section 2.4 that allows correlation between cross-price derivatives and either the tax error or the efficiency rating. We show that when calibrating to estimates of this correlation from the literature the  $R^2$  remains very close to the true fraction of welfare recovered.

To evaluate the level of deadweight loss—following the formula in Equation 6—we first assign a value of \$40 for the social cost of carbon (Interagency Working Group on Social Cost of Carbon (2013)), leading to an external cost of 35.5 cents per gallon.<sup>33</sup> Using our data on lifetime fuel use this implies an average of \$3,334 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 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 these assumptions on elasticities the welfare gain from a Pigouvian tax on each of 356 vehicle

---

<sup>33</sup>If the cost associated with carbon emissions has been rising approximately at the discount rate, we interpret this value as being in 2013 dollars (looking retrospectively at the 1988-1992 vintages).

models amounts to \$246 per car sold, or about \$3.5 billion, for model year 1990. The best linear tax on fuel use per mile, equivalent here to the optimal average fuel-economy standard, generates about \$0.8 billion in surplus and so leaves \$2.7 billion in deadweight loss. This corresponds directly to the intuition on  $R^2$  above: for the 1990 model year the weighted  $R^2$  is 0.24, implying 24% of possible gains can be recovered with a single linear policy.

Table 3 presents the central case described above followed by three panels exploring sensitivity to own- and cross-price elasticities. Panel 1 considers changes in the own-price elasticity of demand for individual vehicle 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 Pigouvian benchmark. As expected the ratio of welfare gains in the second-best vs. the Pigouvian benchmark remains fixed at 0.24, the value of  $R^2$ .

Panel 2 turns to relaxing Assumption 3, investigating how different correlations between cross-price elasticities and the residuals in the policy regressions influence the share of welfare recovered by second-best policy. We first consider the expected direction of bias building on the theory in Section 2.4. Specifically, we demonstrate that  $R^2$  is biased upwards or downwards in this case as predicted by Proposition 3. The first row in Panel 2 makes vehicles with similar durability better substitutes, leaving all other attributes uncorrelated. As expected, this reduces the effectiveness of the Pigouvian benchmark policy and the true fraction of welfare gained increases relative to  $R^2$ . Here, when elasticities fall in half for each standard deviation difference in durability, we see the fraction of welfare recovered increases to 0.27. The second row examines correlation between cross effects and fuel-economy rating, now making cars with similar miles-per-gallon (MPG) better substitutes. This reduces welfare gains in both the Pigouvian benchmark and the second-best; the fraction of welfare recovered falls to 0.19. Finally, the third experiment makes vehicles of similar price the best substitutes. This introduces both types of correlation together since both MPG and durability are related to price. The effects on the fraction of welfare recovered partially offset, with the fraction recovered returning toward  $R^2$ . This further supports the use of the  $R^2$  measure.

Finally, in Panel 3, we calibrate cross-price derivatives using estimates of brand and class loyalty from the literature. This introduces a whole range of correlations together, with class loyalty looking most like correlation with MPG, and brand loyalty tending to create correlation with durability. The first line of this panel shows the net effects in our calculation when calibrating to the brand and class loyalty estimates from the demand system in Bento, Goulder, Jacobsen, and von Haefen (2009). The various effects offset almost completely, with the fraction of welfare recovered falling slightly to 0.23. The final row doubles the strength of the effects in Bento et al. (2009) (doubling the fraction of buyers who substitute within brand and class) and again the effects are very close to offsetting. Across a wide range of substitution patterns,  $R^2$  remains a robust predictor for the fraction of welfare that can be recovered in the second-best.

**Table 3:** Welfare Effects for Model Year 1990

	Second-best	Pigouvian benchmark	Ratio
<b>Central case</b>	817	3472	0.24
<b>Panel 1: Own-price elasticity</b>			
-3	501	2128	0.24
-7	1126	4782	0.24
<b>Panel 2: Cross-price elasticities correlated with:</b>			
Durability	811	2971	0.27
Efficiency rating	638	3385	0.19
Price	739	3523	0.21
<b>Panel 3: Brand and class loyalty</b>			
Calibrated to Bento et al. (2009)	783	3431	0.23
Doubling relative to Bento et al.	756	3396	0.22

Notes: Welfare gains are expressed in millions of 2013 dollars relative to a constant tax at the average externality. For Panel 2, each standard deviation reduction in attribute distance increases the cross-price elasticity by a factor of two.

## 6 Application 4: Spatial Variation in Emissions

Externalities from pollution are typically a function of the amount of pollution emitted as well as location-specific conditions, including pre-existing pollution levels, weather, and the proximity of vulnerable populations. This has long been understood as a rationale for location-specific environmental policies (Tietenberg 1980; Mendelsohn 1986; Baumol and Oates 1988). But, many policies are constrained to be uniform across space, due to practical or political considerations.

When our assumptions hold, our model shows that second-best environmental policies that are constrained to be uniform across space can be analyzed via regression statistics. Many past studies in this area have implicitly assumed our stronger condition, Assumption 3(b'), by assuming that the price of pollution in one location has no impact on the demand for pollution in other locations.<sup>34</sup> For example, Mendelsohn (1986) provides a calibrated example showing the welfare gains from drawing two or three distinct zones that have differentiated policies for air pollution. The exercise assumes that there are no cross-jurisdiction relationships in emissions quantities. This is a special case of our more general  $R^2$  result, in which the explanatory variables that determine the second-best tax scheme are dummy variables for geographic zones.

For another example, Holland et al. (2016) document heterogeneity in the environmental benefit of switching from a gasoline vehicle to an electric vehicle in each U.S. county. They then describe the welfare benefits of a fully differentiated (county level) policy, versus a national policy, versus an in-between policy that varies by state. They assume no demand spillovers across markets, so

<sup>34</sup>Other strains of the voluminous literature on spatial differentiation have focused on emissions leakage (see, e.g., Felder and Rutherford 1993; Paltsev 2001), which can be thought of as cross-price derivatives in our setting. Our regression results may hold in those cases, depending on the pattern of cross-price derivatives.

our model suggests that the fraction of the gains from complete differentiation (as compared to a national policy) achieved by a state-level policy would be the  $R^2$  of the county-level damages on dummy variables for state, with regression weights to account for differences in market sizes.

As such, our core theory model provides a unifying framework that nests some prior literature concerned with spatial variation. To that literature, we introduce the welfare interpretation of regression statistics, which could be used to summarize the relative efficiency of many alternative policies.

In our final application, we consider a variant on geographic differentiation in which there are many products with different externalities within each jurisdiction. Specifically, we consider a tax placed on the purchase of new refrigerators that aims to correct for greenhouse gas emissions. Refrigerators vary in their energy consumption, and the externality of a given refrigerator depends on its location of use, because a unit of electricity corresponds to different amounts of greenhouse gas emissions in different locations according to what type of power is used (e.g., coal versus renewables). We consider a constrained policy under which the tax scheme on appliances depends on only the appliance’s energy consumption, not on its location.

We show that our assumptions will not hold in this case and that the  $R^2$  will be significantly biased. Instead, we offer an alternative derivation that demonstrates that the within- $R^2$  from a fixed effects regression has a welfare interpretation. The main purpose of this exercise is to demonstrate the promise of adapting our core framework to find alternative relationships between familiar regression statistics and the welfare properties of constrained externality-correcting policies.

## 6.1 The Pigouvian benchmark versus the constrained policy

To model spatial variation, we suppose that there are  $s = 1, \dots, S$  geographically distinct markets, and denote product  $j$  sold in market  $s$  as a product  $x_{js}$ . We model the carbon externality from a refrigerator as follows. First, each product has an energy efficiency rating, which is measured in kWh per year (more on this below). To translate energy consumed into carbon emissions, this is multiplied by  $r_s$ , which is the carbon emissions rate per kWh; this varies across electricity markets depending on the marginal source of electricity. This yields a measure of carbon emissions per year, which is then multiplied by a constant  $\omega$ , which scales annual emissions into lifetime emissions. We know of no information about the heterogeneity of lifetimes for different types of refrigerators, so we abstract from the longevity considerations that we explored in our previous application and assume all refrigerators have a common lifetime. Given that assumption, we just use annual energy consumption to calculate emissions rates and tax rates, as scaling them by a lifetime utilization term will have no impact on the  $R^2$ .

The policy we consider is a national policy that creates a tax schedule that depends linearly on energy efficiency rates, but does not differentiate by location:

$$\begin{aligned}\phi_{js} &= \omega \times r_s \times \text{Energy Efficiency}_j \\ t_{js} &= t_j = \alpha + \beta \times \text{Energy Efficiency}_j.\end{aligned}$$

The U.S. does not have such a tax policy for refrigerators, but we believe this is a useful characterization of various existing policies that do treat appliances according to energy efficiency but not location. For example, Energy Star certification for appliances and buildings, which determines eligibility for a variety of subsidies and rebates, depends solely on energy efficiency, not on location. Tax credits for alternative fuel vehicles, weatherization, or solar panels have a similar structure. 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.

## 6.2 Will cross effects simplify?

Assumption 3 will not hold in this case and the cross effects will not simplify. Denote the energy efficiency of product  $j$  as  $f_j$ , and let  $\alpha^*$  and  $\beta^*$  be the second-best tax parameters. Then,  $e_{js} = \phi_{js} - t_{js} = \omega r_s f_j - (\alpha^* + \beta^* f_j) = (\omega r_s - \beta^*) f_j - \alpha^*$ . This shows that the tax errors have a market-specific component,  $\omega r_s - \beta^*$ , which will be shared by all products in a market  $s$ . If we assume that most substitution occurs within a market, instead of across a geographic border (which is clearly a good assumption for large appliances), then the cross-price derivatives will be systematically larger for products within a market  $s$ . This means that there is a systematic correlation between the product  $e_{js} e_{ks}$  and the cross-price derivatives  $\partial x_{js} / \partial t_{ks}$  within each region  $s$ . In brief, in our setting errors in policy will be highly correlated within region, and products sold in the same market are necessarily closer substitutes. Thus, we next derive an alternative result that does apply to our setting.<sup>35</sup>

## 6.3 An alternative: the within- $R^2$ as a welfare measure

To proceed, we make assumptions about the nature of the geographic markets. The first is to assume that consumers reside and use the  $j$  products within a single geographic market. This is stated formally in Assumption 4:

**Assumption 4.** *Geographic markets are separable:  $\frac{\partial x_{js}}{\partial t_{kq}} = 0 \forall j, k$  whenever  $q \neq s$ .*

Under a hypothetical Pigouvian benchmark policy where tax rates differ by region this assumption would also imply that cross-border shopping and resale are not possible. While certain taxes—for example taxes on automobiles—are assessed based on state of residence rather than state of purchase, difficulty in enforcing differing tax rates across geography is one reason the second-best policy we study here (a single tax profile across all regions) may be more likely in practice.

Next, for expositional simplicity, we assume that the  $S$  markets have identical demand systems, so that cross-price derivatives and total demand (conditional on prices) are the same in each market (Assumption 5). Finally, we also make use of an assumption that says that the overall size of each product market is fixed (Assumption 6).

---

<sup>35</sup>In this case, the second-best constrained policy will turn out to nevertheless be the OLS policy, but the  $R^2$  will be a biased measure of its efficiency. In other settings, the second-best policy will differ from the OLS prediction.

**Assumption 5.** Demand is identical in each geographic market:  $x_{js}(t) = x_{jq}(t) \forall t, q, s$ .

**Assumption 6.** Total demand in each geographic market is fixed:  $\sum_j x_{js}(t) = \bar{x}_s \forall s, t$ .

Assumption 5 simplifies notation greatly, but is straightforward to relax and a weighted least-squares intuition (with weights based on the size of regions) applies as in Proposition 2. Assumption 6 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—Assumption 6 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 small, which makes this assumption appealing. That is, virtually every home has a refrigerator, and most have exactly one. Below we demonstrate that welfare results can be bounded when relaxing this assumption.

Under Assumption 4, the deadweight loss from imposing a tax vector  $t = \tau$  can be written as our original formula in Equation 8 summed over the  $S$  regions:

$$-2\text{DWL} = \sum_{s=1}^S \left( \sum_{j=1}^J e_{js}^2 \frac{\partial x_{js}}{\partial t_{js}} + \sum_{j=1}^J \sum_{k \neq j} e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \right). \quad (16)$$

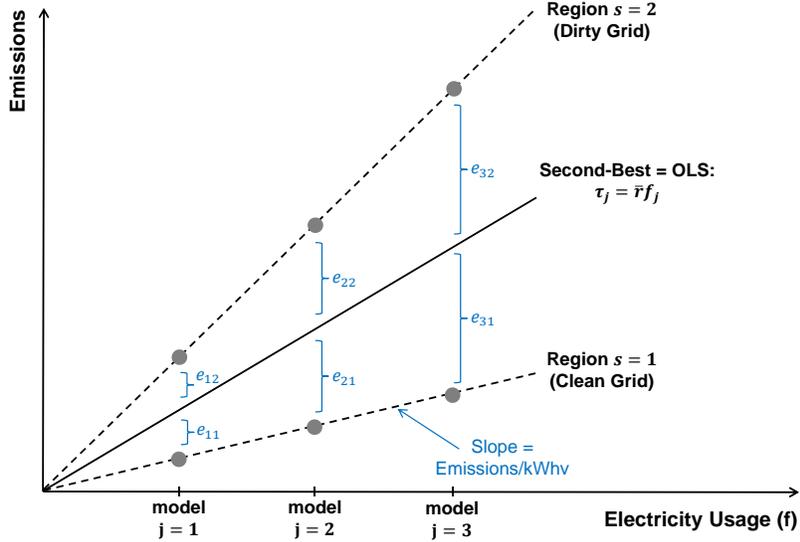
Furthermore, under Assumptions 4 to 6, the second-best national tax policy will be to set the policy slope equal to the average damage factor across regions ( $\beta = \bar{r}$ ).<sup>36</sup> The resulting tax error for a product will be the difference between the local damage factor and the average, multiplied by the attribute:  $e_{js} = r_s f_j - t_{js} = (r_s - \bar{r}) f_j$ . Within each region, these second-best taxes create two types of mis-pricing. First, prices are biased (compared to the Pigouvian benchmark), 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 underpriced, but so are its substitutes (the other products in region 2). This implies that cross effects will mitigate own effects (the second term of Equation 16 partially offsets the first) and the raw  $R^2$  statistic in Corollary 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 either 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 tax errors.

<sup>36</sup>A proof of this is included within the proof of Proposition 4 in Appendix A.

**Figure 5:** Average and Relative Mis-Pricing across Regions



Assumption 6 implies that the first type of mis-pricing—average bias within a region—creates no distortion in choice.<sup>37</sup> In this case, a simple sufficient statistic from a fixed effects regression 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 4 (see Appendix A for a proof).

**Proposition 4.** *Under Assumptions 1, 2, 4, 5 and 6, the second-best policy is  $t_j = \bar{r}f_j$ . The fraction of the welfare gain from the Pigouvian benchmark achieved by this second-best policy is the within- $R^2$  from a regression of  $\phi_{js}$  on  $f_j$  with fixed effects for the  $S$  regions:*

$$\frac{DWL(t = \bar{r}f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})} = 1 - \frac{\text{var}(r_s)}{E[r_s^2]} = \text{within-}R^2. \quad (17)$$

As with our results in Section 2 this is a familiar statistic from regression analysis: The within- $R^2$  from a fixed effects regression is provided automatically by most statistical software or can be readily calculated as the  $R^2$  from an OLS regression that has been de-meanned by region. In our setting this statistic depends only on variation in damage factors across regions (the first equality in Equation 17), meaning it again does not rely on any information about the demand system (e.g., the second-type of mis-pricing mentioned above is irrelevant) beyond the assumptions on geographic separability and net substitution to the outside good. The structure of demand within each region will influence the deadweight loss measured in dollars,<sup>38</sup> but variation in the demand system affects welfare under the policy alternatives in a proportional way such that all the demand

<sup>37</sup>The constant  $\alpha$  in a tax schedule  $t_{js} = \alpha + \beta f_j$  is therefore irrelevant.  $\alpha$  does affect a transfer between consumers and the government, but this is undone through revenue recycling.

<sup>38</sup>As shown in the proof of Proposition 4 in Appendix A, the deadweight loss expression can be obtained by

derivatives divide through.

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 increases expand the market). Under that assumption, the welfare statistic derived above will overstate the fractional welfare gain of the second-best policy over the baseline of an equal tax on all products, which we state in Corollary 2.<sup>39</sup>

**Corollary 2.** *Under Assumptions 1, 2, 4, and 5, the second-best policy slope remains  $\beta = \bar{r}$ , and the fraction of the welfare gain from the Pigouvian benchmark achieved by this second-best policy over a policy of a constant tax on all products is:*

$$\frac{DWL(t = \bar{r}f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})} < \text{within-}R^2. \quad (18)$$

Thus, the same sufficient statistic now represents an upper bound on the fraction of the welfare gain under the Pigouvian benchmark that is achieved by the non-spatially differentiated product tax. Intuitively, the second-best policy will not have the correct slope in all regions (as before) and now it also over- or under-taxes with respect to the outside good. In our empirical application, which we move to next, substitution to an outside good is probably of limited importance, but this effect could be important in other applications.

## 6.4 Data

Implementation of our method requires data on energy efficiency rates of different refrigerator models and carbon emissions rates per unit of electricity for each region of the country. 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. 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.

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

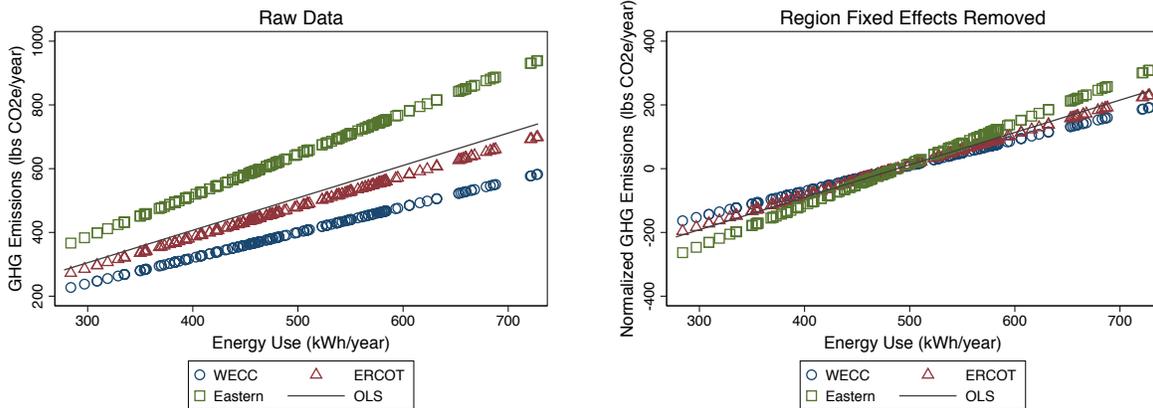
---

simplifying the expression in Equation 16:

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

<sup>39</sup>The proof is in Appendix A. For expositional clarity, the proof uses an additional regularity condition stating that average substitution to the outside good across products is not correlated with the attribute or 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 Western Electricity Coordinating Council (WECC), the Electric Reliability Council of Texas (ERCOT) and the Eastern Interconnection (Eastern).

2014; Holland et al. 2016).

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

## 6.5 Results

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 left panel 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 4. 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 data with region fixed effects removed. Visually, the remaining variation determines the within- $R^2$ . 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 4 shows the  $R^2$  and within- $R^2$  values from a regression of the carbon emissions associated with each product-region level observation on the product’s electricity consumption rate. The first row includes region fixed effects that produce within- $R^2$  values of 0.96 and 0.90 (depending on geographic disaggregation), representing the fraction of the welfare gain under the Pigouvian benchmark achieved by a national policy that does no spatial differentiation. For reference, the  $R^2$  values from OLS, which are far smaller, are also included.

These findings demonstrate a, perhaps surprisingly, small welfare loss from the lack of regional policy differentiation for electric appliances. The estimates from OLS 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 mispricing of products *within* a market are highly correlated, muting the degree to which mispricing 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 Pigouvian benchmark policy over the baseline). This result relies on the extensive margin for refrigerator demand being zero or small. When the extensive margin response grows (and it will be larger for appliances other 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.

## 7 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, technological cost, or political constraints so that they must place imperfect marginal incentives on products or actions. We demonstrate that, under certain conditions, simple regression statistics have welfare interpretations that describe the efficiency costs of these constraints.

We demonstrate the usefulness of this approach through four examples. Three of our applications pertain to environmental externalities, but they span a number of distinct challenges to policy, including random mismeasurement of product attributes, spatial heterogeneity, and the

**Table 4:** Sufficient Statistics for Spatially Differentiated Refrigerators

	<b>Three interconnections</b>	<b>Eight NERC regions</b>
Within- $R^2$ , fixed effects	0.96	0.90
$R^2$ , OLS	0.47	0.24
Sample size	4,047	10,792

Note: Results are from regressing emissions on electricity consumption with the unit of observation a refrigerator model in a particular interconnection (second column) or NERC region (third column).

implications of heterogeneity in the lifetime utilization of energy-consuming durable goods. Our other application, which is about wedges between price and marginal cost due to coarse pricing rather than externalities, suggests the potentially wider reach of our approach. These applications 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.

We suspect that there are many additional applications that could benefit from this approach. In the introduction, we mention other possible applications in energy, environment, health and transportation, but the possibilities extend to any setting where data are available on the distribution of an externality (or other wedge) and its correlation with the variables upon which policy is contingent. Some of our results may be relevant to settings where there is heterogeneity in the deadweight loss of taxation even in the absence of externalities. For example, labor supply elasticities differ along dimensions such as age—the young supply labor more inelastically than the old (Kleven and Schultz 2014).<sup>40</sup> It is generally politically infeasible to condition income or payroll taxes on age. Our findings suggest that these restrictions, while perhaps desirable based on other grounds, increase the overall deadweight loss of labor taxation and provide a method to quantify the efficiency loss. In applying our model to other settings, we emphasize that it is important to consider the demand assumptions, but also note that it is straightforward to conduct robustness checks that indicate the degree of error created when the assumptions do not hold.

## References

- Allcott, Hunt and Michael Greenstone. 2012. “Is There an Energy Efficiency Gap?” *Journal of Economic Perspectives* 26 (1):3–28.
- Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky. 2014. “Energy Policy with Externalities and Internalities.” *Journal of Public Economics* 112:72–88.
- Andersen, Laura M., Lars G. Hansen, Carsten L. Jensen, and Frank A. Wolak. 2017. “Using Real-Time Pricing and Information Provision to Shift Intra-Day Electricity Consumption: Evidence from Denmark.” URL [http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/into\\_versus\\_away\\_paper.pdf](http://web.stanford.edu/group/fwolak/cgi-bin/sites/default/files/into_versus_away_paper.pdf). Manuscript: Stanford University.
- Anderson, Soren T., Ian W.H. Parry, James M. Sallee, and Carolyn Fischer. 2011. “Automobile Fuel

---

<sup>40</sup>Best and Kleven (2013) show theoretically that the presence of behavioral career effects provides another reason why the contemporaneous earnings elasticity of the young is lower than of the old.

- Economy Standards: Impacts, Efficiency and Alternatives.” *Review of Environmental Economics and Policy* 5 (1):89–108.
- Anderson, Soren T. and James M. Sallee. 2016. “Designing Policies to Make Cars Greener: A Review of the Literature.” *Annual Review of Resource Economics* 8:157–180.
- Ata, Baris, A. Serasu Duran, and Ozge Islegen. 2016. “An Analysis of Time-Based Pricing in Electricity Supply Chains.” URL <https://ssrn.com/abstract=2826055>. Manuscript: University of Chicago.
- Baily, Martin N. 1978. “Some Aspects of Optimal Unemployment Insurance.” *Journal of Public Economics* 10 (3):379–402.
- Baumol, William J. and Wallace E. Oates. 1988. *The Theory of Environmental Policy*. Cambridge University Press, 2nd ed.
- Bento, Antonio M., Lawrence H. Goulder, Mark R. Jacobsen, and Roger H. von Haefen. 2009. “Distributional and Efficiency Impacts of Increased US Gasoline Taxes.” *American Economic Review* 99 (3):667–699.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica* 64 (4):841–890.
- Best, Michael C. and Henrik J. Kleven. 2013. “Optimal Income Taxation with Career Effects of Work Effort.” URL [http://www.henrikkleven.com/uploads/3/7/3/1/37310663/best-kleven\\_careers\\_feb2013.pdf](http://www.henrikkleven.com/uploads/3/7/3/1/37310663/best-kleven_careers_feb2013.pdf). Manuscript: London School of Economics.
- Borenstein, Severin. 2005. “The Long-Run Efficiency of Real-Time Electricity Pricing.” *Energy Journal* 26 (3):93–116.
- Borenstein, Severin and Stephen P. Holland. 2005. “On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices.” *RAND Journal of Economics* 36 (3):469–493.
- Callaway, Duncan S., Meredith Fowlie, and Gavin McCormick. 2018. “Location, Location, Location: The Variable Value of Renewable Energy and Demand-Side Efficiency Resources.” *Journal of the Association of Environmental and Resource Economists* 5 (1):39–75.
- Caltrans. 2016. “Performance Measurement System.” URL <http://pems.dot.ca.gov/>.
- Chetty, Raj. 2006. “A General Formula for the Optimal Level of Social Insurance.” *Journal of Public Economics* 90 (10-11):1879–1901.
- . 2009. “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods.” *Annual Review of Economics* 1 (1):451–488.
- Chetty, Raj, Adam Looney, and Kory Kroft. 2009. “Salience and Taxation: Theory and Evidence.” *American Economic Review* 99 (4):1145–1177.
- Cullen, Joseph A. 2013. “Measuring the Environmental Benefits of Wind Generated Electricity.” *American Economic Journal: Economic Policy* 5 (4):107–133.
- . 2015. “Dynamic Response to Environmental Regulation in Electricity Markets.” URL <http://www.josephcullen.com/resources/dynamicresponse.pdf>. Manuscript: Washington University.

- 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.
- Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. “Estimating Welfare in Insurance Markets Using Variation in Prices.” *The Quarterly Journal of Economics* 125 (3):877–921.
- Farrow, R. Scott, Martin T. Schultz, Pinar Celikkol, and George L. van Houten. 2005. “Pollution Trading in Water Quality Limited Areas: Use of Benefits Assessment and Cost-Effective Trading Ratios.” *Land Economics* 81 (2):191–205.
- Felder, S. and T.F. Rutherford. 1993. “Unilateral CO<sub>2</sub> Reductions and Carbon Leakage: The Effect of International Trade in Oil and Basic Materials.” *Journal of Environmental Economics and Management* 25 (2):162–176.
- Feldstein, Martin. 1999. “Tax Avoidance and the Deadweight Loss of the Income Tax.” *Review of Economics and Statistics* 81 (4):674–680.
- Feng, Ye, Don Fullerton, and Li Gan. 2013. “Vehicle Choices, Miles Driven, and Pollution Policies.” *Journal of Regulatory Economics* 44 (1):4–29.
- Fowlie, Meredith and Nicholas Z. Muller. 2017. “Market-based Emission Regulation when Damages Vary Across Sources: What Are the Gains from Differentiation?” URL [https://www.meredithfowlie.com/s/Fowlie\\_Muller\\_2017.pdf](https://www.meredithfowlie.com/s/Fowlie_Muller_2017.pdf). Manuscript: UC Berkeley.
- Fowlie, Meredith, Catherine Wolfram, C. Anna Spurlock, Annika Todd, Patrick Baylis, and Peter Cappers. 2017. “Default Effects and Follow-on Behavior: Evidence from an Electricity Pricing Program.” URL <https://www.meredithfowlie.com/s/Default-Effects-SMUD.pdf>. 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.
- Gillan, James. 2017. “Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption.” URL [http://jamesgillan.info/s/JMP\\_Gillan\\_most\\_recent\\_draft.pdf](http://jamesgillan.info/s/JMP_Gillan_most_recent_draft.pdf). Manuscript: UC Berkeley.
- Gillitzer, Christian, Henrik J. Kleven, and Joel Slemrod. 2017. “A Characteristics Approach to Optimal Taxation: Line Drawing and Tax-Driven Product Innovation.” *Scandinavian Journal of Economics* 119 (2):240–267.
- Goulder, Lawrence H. and Roberton C. Williams. 2003. “The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution.” *Journal of Political Economy* 111 (4):898–927.
- 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.

- Harrington, Winston, Ian Parry, and Margaret Walls. 2007. “Automobile Externalities and Policies.” *Journal of Economic Literature* 45 (2):373–399.
- Hendren, Nathaniel. 2016. “The Policy Elasticity.” *Tax Policy and the Economy* 30 (1):51–89.
- 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. 2016. “Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors.” *American Economic Review* 106 (12):3700–3729.
- Interagency Working Group on Social Cost of Carbon. 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>.
- Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. “Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand.” *American Economic Journal: Economic Policy* 10 (1):1–29.
- Ito, Koichiro and James M. Sallee. Forthcoming. “The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards.” *Review of Economics and Statistics* .
- Jacobsen, Mark R. and Arthur A. van Benthem. 2015. “Vehicle Scrappage and Gasoline Policy.” *American Economic Review* 105 (3):1312–1338.
- Jessoe, Katrina and David Rapson. 2014. “Knowledge is (Less) Power: Experimental Evidence from Residential Electricity Use.” *American Economic Review* 104 (4):1417–1438.
- Kleven, Henrik J. and Claus Thustrup Kreiner. 2006. “The Marginal Cost of Public Funds: Hours of Work versus Labor Force Participation.” *Journal of Public Economics* 90 (10-11):1955–1973.
- Kleven, Henrik J. and Esben A. Schultz. 2014. “Estimating Taxable Income Responses Using Danish Tax Reforms.” *American Economic Journal: Economic Policy* 6 (4):271–301.
- 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.” URL [http://www.sdmcrae.com/wp-content/uploads/var\\_fuel\\_econ\\_langer\\_mcrae.pdf](http://www.sdmcrae.com/wp-content/uploads/var_fuel_econ_langer_mcrae.pdf). Manuscript: University of Arizona.
- Leard, Benjamin and Virginia McConnell. Forthcoming. “New Markets for Pollution and Energy Efficiency: Credit Trading Under Automobile Greenhouse Gas and Fuel Economy Standards.” *Review of Environmental Economics and Policy* .
- Majeske, Karl D., Terri Lynch-Caris, and Janet Brelin-Fornari. 2010. “Quantifying R2 Bias in the Presence of Measurement Error.” *Journal of Applied Statistics* 37 (4):667–677.
- Mendelsohn, Robert. 1986. “Regulating Heterogeneous Emissions.” *Journal of Environmental Economics and Management* 13 (4):301–312.

- Montero, Juan-Pablo. 1999. "Voluntary Compliance with Market-Based Environmental Policy: Evidence from the U. S. Acid Rain Program." *Journal of Political Economy* 107 (5):1145–1177.
- 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 the United States Economy." *American Economic Review* 101 (5):1649–1675.
- National Research Council. 2010. *Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use*. Washington, DC: National Academy Press.
- Paltsev, S.V. 2001. "The Kyoto Protocol: Regional and Sectoral Contributions to the Carbon Leakage." *The Energy Journal* 22 (4):53–79.
- Pigou, Arthur C. 1932. *The Economics of Welfare*. London: Macmillan and Co., 4th ed.
- Reguant, Mar. 2014. "Complementary Bidding Mechanisms and Startup Costs in Electricity Markets." *Review of Economic Studies* 4 (81):1708–1742.
- Saez, Emmanuel, Joel B. Slemrod, and Seth H. Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50 (1):3–50.
- Sallee, James M. 2014. "Rational Inattention and Energy Efficiency." *Journal of Law and Economics* 57 (3):781–820.
- Tietenberg, Tom H. 1980. "Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis." *Land Economics* 56 (4):391–416.
- Urban, Lorien E., Megan A. McCrory, Gerard E. Dallal, Sai Krupa Das, Edward Saltzman, Judith L. Weber, and Susan B. Roberts. 2011. "Accuracy of Stated Energy Contents of Restaurant Foods." *The Journal of the American Medical Association* 306 (3):287–293.

## A Appendix: Proofs

### Derivation of Equation 6

Let any generic tax schedule be denoted as  $\tau_1, \dots, \tau_J$ . To obtain Equation 6, we characterize the welfare loss of moving from the Pigouvian benchmark tax schedule  $t_j = \phi_j$  to  $t_j = \tau_j$  by specifying a weighted average of the two tax schedules and then integrating the marginal welfare losses of moving the weights from  $\phi_j$  to  $\tau_j$ . That is, we specify the function  $t_j = (1 - \rho)\phi_j + \rho\tau_j$ . We differentiate  $W$  with respect to  $\rho$ , and then derive the welfare loss of moving from the Pigouvian policy (when  $\rho = 0$ ) to the alternative policy (when  $\rho = 1$ ).

First, we differentiate Equation 2 with respect to  $\rho$  and substitute in the consumer's optimality condition. This yields:

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

This term,  $\frac{dW}{d\rho}$ , is the incremental change in welfare as we move from the Pigouvian benchmark rates toward the alternative tax schedule, where all rates move by an amount proportional to the difference between the Pigouvian benchmark taxes 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:

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

Because  $\rho$  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:

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

### Allowing for increasing marginal cost

In our theoretical derivation, we assume a constant marginal cost. This means that the incidence of a tax is borne completely by consumers ( $\partial p / \partial t = 1$ ), and in turn that  $\partial x_j / \partial t_k$  is interpreted

directly as a demand derivative. Our mathematical derivations remain the same if we allow for increasing, but locally linear, marginal costs. The interpretation of the assumptions, however, are now in terms of equilibrium quantity changes  $\frac{\partial \tilde{x}_j}{\partial t_k}$ . These relate to combined supply and demand derivatives, rather than just demand derivatives. This is equivalent to adding assumptions about pass through.

With increasing marginal cost, Assumption 3(b) requires that  $cov\left(e_j e_k, \frac{\partial \tilde{x}_j}{\partial t_j}\right) = 0$ . This assumption is met if the errors are orthogonal to equilibrium quantity responses to cross-tax rate changes. This substitution affects the interpretation of Assumption 3, which is a statement about whether the *variation* in demand derivatives is correlated with functions of the tax errors. Introducing increasing marginal cost that is uniform across products would have no material impact on the interpretation of our assumptions. Neither would cost functions for which the quantity of good  $j$  produced does not affect the production cost of good  $k \neq j$ . Specifically, denote demand as  $d_j$  and supply as  $s_j$ . When cross-price demand derivatives  $\frac{\partial d_j}{\partial t_k} = 0$  (Assumption 3(b')) and cross-price supply derivatives  $\frac{\partial s_j}{\partial t_k} = 0$ ,  $\frac{\partial \tilde{x}_j}{\partial t_j} = 0$  and all our results from Section 2 go through (these assumptions are sufficient but not necessary).

## Derivation of Equations 9 and 10

Begin with the general welfare formula:

$$DWL(t = \tau) = -\frac{1}{2} \sum_{j=1}^J e_j^2 \frac{\partial x_j}{\partial t_j} - \frac{1}{2} \sum_{j=1}^J \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k} \quad (\text{A.4})$$

Applying the part (a) of Assumption 3 we can move the average own-price derivative out from the own effects:

$$= -\frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} \sum_{j=1}^J e_j^2 - \frac{1}{2} \sum_{j=1}^J \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k} \quad (\text{A.5})$$

When the strong version (b') of Assumption 3 holds, (A.5) directly reduces to Equation 9.

Substituting in the weaker version (b) of Assumption 3 (i.e.,  $cov\left(e_j e_k, \frac{\partial x_j}{\partial t_j}\right) = 0$ ), we obtain:

$$= -\frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} \sum_{j=1}^J e_j^2 - \frac{1}{2} \frac{1}{J(J-1)} \left( \sum_{j=1}^J \sum_{k \neq j} e_j e_k \right) \left( \sum_{j=1}^J \sum_{k \neq j} \frac{\partial x_j}{\partial t_j} \right) \quad (\text{A.6})$$

Using  $\sum_{k \neq j} e_k = \sum_{k=1}^J e_k - e_j$ , the expression becomes:

$$= -\frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} \sum_{j=1}^J e_j^2 - \frac{1}{2} \frac{1}{J(J-1)} \left( \left( \sum_{j=1}^J e_j \right)^2 - \sum_{j=1}^J e_j^2 \right) \left( \sum_{j=1}^J \sum_{k \neq j} \frac{\partial x_j}{\partial t_j} \right) \quad (\text{A.7})$$

Now impose the mean-zero error condition to obtain an expression equivalent to Equation 10 in the main text:

$$\begin{aligned}
&= -\frac{1}{2} \left( \overline{\frac{\partial x_j}{\partial t_j}} - \frac{1}{J(J-1)} \left( \sum_{j=1}^J \sum_{k \neq j} \frac{\partial x_j}{\partial t_k} \right) \right) \sum_{j=1}^J e_j^2 \\
&= -\frac{1}{2} \left( \overline{\frac{\partial x_j}{\partial t_j}} - \overline{\frac{\partial x_j}{\partial t_k}} \right) \sum_{j=1}^J e_j^2.
\end{aligned} \tag{A.8}$$

### Proof of Proposition 1

We apply Assumptions 1 through 3 to the tax errors and minimize the deadweight loss formula. The first steps follow the derivation of Equation 10 above, beginning with the general welfare formula in (A.4). We show the derivation for a single policy variable  $f_j$ , but the proof trivially extends to multivariate case. We do not yet impose an unbiased tax—this will be endogenous to the choice of  $\alpha$  and  $\beta$ —so the relevant objective function is in the same form as (A.7):

$$\begin{aligned}
\min_{\alpha, \beta} DWL(t = \tau) &= -\frac{1}{2} \sum_{j=1}^J \sum_{k=1}^J (\phi_j - \alpha - \beta f_j)(\phi_k - \alpha - \beta f_k) \frac{\partial x_j}{\partial t_k} \\
&= -\frac{1}{2} \left( \overline{\frac{\partial x_j}{\partial t_j}} - \overline{\frac{\partial x_j}{\partial t_k}} \right) \sum_{j=1}^J (\phi_j - \alpha - \beta f_j)^2 - \frac{1}{2} \overline{\frac{\partial x_j}{\partial t_k}} \left( \sum_{j=1}^J (\phi_j - \alpha - \beta f_j) \right)^2
\end{aligned} \tag{A.9}$$

Noting that the average derivatives are constant with respect to the policy choice, we take first-order conditions:

$$\begin{aligned}
\alpha : \overline{\frac{\partial x_j}{\partial t_j}} \left( \sum_{j=1}^J \phi_j - \alpha J - \beta \sum_{j=1}^J f_j \right) &= 0 \\
\beta : \left( \overline{\frac{\partial x_j}{\partial t_j}} - \overline{\frac{\partial x_j}{\partial t_k}} \right) \left( \sum_{j=1}^J f_j \phi_j - \alpha \sum_{j=1}^J f_j - \beta \sum_{j=1}^J f_j^2 \right) &+ \overline{\frac{\partial x_j}{\partial t_k}} \left( \sum_{j=1}^J f_j \right) \left( \sum_{j=1}^J \phi_j - \alpha J - \beta \sum_{j=1}^J f_j \right) = 0
\end{aligned}$$

Solving:

$$\begin{aligned}
\beta &= \frac{\sum_{j=1}^J f_j \phi_j - \frac{1}{J} \sum_{j=1}^J f_j \sum_{j=1}^J \phi_j}{\sum_{j=1}^J f_j^2 - \frac{1}{J} (\sum_{j=1}^J f_j)^2} = \beta^{OLS} \\
\alpha &= \frac{1}{J} \sum_{j=1}^J \phi_j - \beta \frac{1}{J} \sum_{j=1}^J f_j = \alpha^{OLS}
\end{aligned} \tag{A.10}$$

The second-order conditions for this problem reduce to:

$$-J \frac{\overline{\partial x_j}}{\partial t_j} > 0$$

$$\left( \frac{\overline{\partial x_j}}{\partial t_j} - \frac{\overline{\partial x_j}}{\partial t_k} \right) \left( \frac{1}{J} \left( \sum_{j=1}^J f_j \right)^2 - \sum_{j=1}^J f_j^2 \right) > 0$$

The residuals  $e$  from an OLS regression become the tax errors and will sum to zero by construction of  $\alpha^{OLS}$  and  $\beta^{OLS}$ . The second term in (A.9) will therefore be zero and deadweight loss at the second-best policy is given by:

$$-\frac{1}{2} \left( \frac{\overline{\partial x_j}}{\partial t_j} - \frac{\overline{\partial x_j}}{\partial t_k} \right) \sum_{j=1}^J e_j^2 = -\frac{1}{2} \left( \frac{\overline{\partial x_j}}{\partial t_j} - \frac{\overline{\partial x_j}}{\partial t_k} \right) SSR \quad (\text{A.11})$$

### Proof of Corollary 1

This follows from the use of Equation 10 and Proposition 1 to evaluate the deadweight losses from the second-best ( $t_j = \alpha^{OLS} + \beta^{OLS} f_j$ ) and uniform ( $t_j = \bar{t} = \bar{\phi}$ ) policies. Define as  $S$  the improvement the second-best tax offers relative to the improvement the Pigouvian benchmark tax offers:

$$S = \frac{DWL(t = \alpha^{OLS} + \beta^{OLS} f_j) - DWL(t = \bar{\phi})}{DWL(t = \phi) - DWL(t = \bar{\phi})} = 1 - \frac{DWL(t = \alpha^{OLS} + \beta^{OLS} f_j)}{DWL(t = \bar{\phi})}$$

where the equality follows from  $DWL(t = \phi) = 0$ .

Define the residuals in the regression of  $\phi$  on a constant (the uniform policy) as  $\lambda_j$ . Note that  $\lambda_j$  has dual significance: it is also the total deviation in the original second-best policy resulting from the regression of  $\phi$  on  $f$ :

$$\sum_{j=1}^J \lambda_j^2 = SSR^{uniform} = TSS^{second-best} \quad (\text{A.12})$$

Finally, define  $\gamma_j$  as the explained portion in the original regression:  $\gamma_j = \alpha^{OLS} + \beta^{OLS} f_j - \bar{\phi}$  so we have  $\lambda_j = \gamma_j + e_j$ . Notice that since  $\gamma$  is a function of  $f$  we have that the tax errors in the uniform policy depend on  $f$ . In order to apply the deadweight loss formula in Equation 10 to the uniform policy we will need to apply Assumption 3: these tax errors must also be uncorrelated with demand derivatives. Under Assumptions 1-3 and applying Equation 10 the fraction  $S$  becomes:

$$S = 1 - \frac{-\frac{1}{2} \left( \frac{\overline{\partial x_j}}{\partial t_j} - \frac{\overline{\partial x_j}}{\partial t_k} \right) SSR^{second-best}}{-\frac{1}{2} \left( \frac{\overline{\partial x_j}}{\partial t_j} - \frac{\overline{\partial x_j}}{\partial t_k} \right) SSR^{uniform}} = 1 - \frac{SSR^{second-best}}{TSS^{second-best}} = R^2 \text{ second-best} \quad (\text{A.13})$$

## Proof of Proposition 2

This follows directly from Proposition 1 and the fact that Equation 7 can be written in matrix notation as:

$$\min_{\mathbf{b}} DWL(\mathbf{t} = \boldsymbol{\tau}) = -\frac{1}{2} \mathbf{e}' \mathbf{D} \mathbf{e} \quad (\text{A.14})$$

where  $\mathbf{D}$  is the matrix of own- and cross-price derivatives of demand, the vector  $\mathbf{e} = \boldsymbol{\phi} - \mathbf{F}\mathbf{b}$  in which  $\boldsymbol{\phi}$  is the vector of product-specific externalities,  $\mathbf{F}$  is the matrix of product attribute values including a constant, and  $\mathbf{b} = (\boldsymbol{\alpha} \boldsymbol{\beta})'$  is the vector of policy coefficients. Now redefine  $\mathbf{D}^* = -\mathbf{D}$ , so that the problem becomes to minimize  $\mathbf{e}' \mathbf{D}^* \mathbf{e}$ . This is exactly the definition of a generalized least squares estimation.<sup>41</sup>

## Proof of Proposition 3

We investigate correlation in cross-price derivatives that would violate Assumption 3, revisiting the proof for Corollary 1. When the assumptions on cross-price derivatives are relaxed the DWL formula in Proposition 1 expands to:

$$DWL = -\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} SSR - \frac{1}{2} \sum_j \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k} \quad (\text{A.15})$$

Equation A.13 in turn expands to:

$$S = 1 - \frac{-\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} SSR^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}{-\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} TSS^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} (\gamma_j + e_j)(\gamma_k + e_k) \frac{\partial x_j}{\partial t_k}} =$$

$$1 - \frac{-\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} SSR^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}}{-\frac{1}{2} \overline{\frac{\partial x_j}{\partial t_j}} TSS^{second-best} - \frac{1}{2} \sum_j \sum_{k \neq j} \gamma_j \gamma_k \frac{\partial x_j}{\partial t_k} - \frac{1}{2} \sum_j \sum_{k \neq j} e_j e_k \frac{\partial x_j}{\partial t_k}} \quad (\text{A.16})$$

The above expression allows cross-price derivatives to be correlated either with  $e$  or with  $\gamma$ . Restricting ourselves to one form of correlation at a time, with no correlation between the correlated pair and the remaining residual term, allows us to remove the terms interacting  $\gamma$  and  $e$  in the denominator. Note that  $\sum_j \sum_{k \neq j} \gamma_j e_k = -\sum_j \gamma_j e_j = 0$  (by construction of the OLS estimate).

Relaxing Assumption 3 to allow correlation between cross-price derivatives and  $\gamma$  (a function of  $f$ ): If cross-price derivatives are large when  $\gamma_j$  and  $\gamma_k$  are similar then the summation involving  $\gamma$  in the denominator of (A.16) becomes larger. Since it is subtracted from the own-price term the fraction grows and  $S$  will decrease relative to the original  $R^2$  measure.

Relaxing Assumption 3 to allow correlation between cross-price derivatives and  $e$ : The same terms involving  $e$  are present in both the numerator and denominator. If cross-price derivatives are

<sup>41</sup>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  $\mathbf{b}^{GLS} = (\mathbf{F}' \mathbf{D}^* \mathbf{F})^{-1} \mathbf{F}' \mathbf{D}^* \boldsymbol{\phi}$ .

large when  $e_j$  and  $e_k$  are similar then the numerator and denominator decline equally. Since the denominator is larger (if there was any information in  $f$ ) the fraction shrinks and  $S$  will increase relative to the original  $R^2$  measure.

## Derivation of the welfare loss statistics for electricity pricing

Here we derive an alternative expression to Equation 9 in the main text for the deadweight loss from using a constrained-optimal tariff structure rather than marginal cost pricing under the real-time pricing benchmark. As discussed in Section 4, we assume zero cross-price derivatives in demand and supply (marginal cost).<sup>42</sup> We then show that the constrained-optimal tariff policy corresponds to a linear regression of the benchmark marginal cost (under real-time pricing) on tariff variables. Finally, we demonstrate how the  $R^2$  from this regression is identical to the  $R^2$  from a regression of the observed marginal cost (under the flat tariff) on the same tariff variables.

## Welfare gain from a regression of *benchmark* marginal cost on tariff variables

We start with the following optimization problem, which is a simplified version from the setup in Section 2:

$$\begin{aligned} \max_{x_1, \dots, x_J} \quad & Z = U(x_1, \dots, x_J) + n \\ \text{s.t.} \quad & \sum_{j=1}^J t_j x_j \leq M + D \\ \text{FOC} \quad & \frac{\partial U}{\partial x_j} = t_j \end{aligned}$$

The consumer consumes up to the point where the marginal utility equals the tariff. The planner's problem has no externalities, but non-constant marginal costs:

$$\max_{t_1, \dots, t_J} W = U(x_1, \dots, x_J) + M - C(x_1, \dots, x_J)$$

With zero cross-price derivatives, we can just write out the DWL from each tax and then sum:

---

<sup>42</sup>Note that we are not assuming that there is no correlation in the cost of production across hours. The cost function will be correlated across adjacent time periods and over parts of the day, for example because renewable generation follows weather patterns. This is not a problem for our model, which makes no assumptions about the correlation of the cost function across products.

$$\begin{aligned}\frac{dW}{dt_j} &= - \left( \frac{\partial U}{\partial x_j} - \frac{\partial C}{\partial x_j} \right) \frac{\partial x_j}{\partial t_j} \\ &= -(t_j - mc_j) \frac{\partial x_j}{\partial t_j}\end{aligned}$$

where  $mc_j$  is the marginal cost. This implies that the real-time pricing benchmark tariff is  $t_j^* = mc_j^* \forall j$ . Note that  $\partial x_j / \partial t_j$  term is the slope of the demand curve. The slope of supply is in  $mc_j$ , not hidden in the  $dx_j / dt_j$  terms. The latter happens when you assume prices are passed through to consumers but add a subsidy/tax to each price at each  $j$ . When demand and supply are locally linear, the deadweight loss depends on the relative slopes of demand and supply as well as the price wedge.

We now show that the deadweight loss expression can be written as a sum of squared errors, where the error is given by the difference between the tariff  $t_j = \tau_j$  and the real-time pricing benchmark tariff  $t_j^* = mc_j^*$ . From that, using the results in Section 2, it follows immediately that the  $R^2$  of a regression of *benchmark* marginal costs on a set of tariff policy variables recovers the welfare gain of the constrained-optimal tariff relative to real-time pricing (with the reference policy being an unbiased flat tariff  $\bar{t}$ ).

We now integrate from  $t_j = mc_j^*$  (the optimum) to some arbitrary  $t_j = \tau_j$ . With no cross effects, we can do this integral for each  $t_j$  and then sum over  $j$ . Let  $mc_j^*$  denote the marginal cost at the efficient tariff and  $mc_j$  the marginal cost for any consumer tariff  $\tau_j$ :

$$\begin{aligned}W(t = t^*) - W(t = \tau) &\equiv DWL(t = \tau) = \sum_{j=1}^J \int_{mc_j^*}^{\tau_j} \frac{dW}{dt_j} dt_j \\ &= \sum_{j=1}^J \int_{mc_j^*}^{\tau_j} -(t_j - mc_j) \frac{\partial x_j}{\partial t_j} dt_j \\ \text{(local linear supply)} &= \sum_{j=1}^J \int_{mc_j^*}^{\tau_j} -(t_j - (mc_j^* + \frac{\partial mc_j}{\partial t_j} (t_j - mc_j^*))) \frac{\partial x_j}{\partial t_j} dt_j \\ \text{(uncorrelated own-derivatives)} &= - \frac{\overline{\partial x_j}}{\partial t_j} \left(1 - \frac{\overline{\partial mc_j}}{\partial t_j}\right) \sum_{j=1}^J \int_{mc_j^*}^{\tau_j} (t_j - mc_j^*) dt_j \\ \text{(anti derivative)} &= - \frac{\overline{\partial x_j}}{\partial t_j} \left(1 - \frac{\overline{\partial mc_j}}{\partial t_j}\right) \sum_{j=1}^J \left[ \frac{1}{2} (t_j - mc_j^*)^2 \right]_{mc_j^*}^{\tau_j} \\ &= - \frac{1}{2} \frac{\overline{\partial x_j}}{\partial t_j} \left(1 - \frac{\overline{\partial mc_j}}{\partial t_j}\right) \sum_{j=1}^J (\tau_j - mc_j^*)^2\end{aligned}$$

Hence, with locally linear demand and marginal costs, the DWL expression is proportional to a sum of squared errors, where the error is defined as the difference between the chosen tariff  $t_j = \tau_j$

and the marginal cost under the real-time pricing benchmark tariff  $t_j = mc_j^*$ . Thus, minimizing DWL is equivalent to minimizing the squared tariff errors, so that the  $R^2$  of a regression of  $mc_j^*$  on the policy variables  $z_j$  indicates the relative efficiency gain:

$$\begin{aligned}
S &= \frac{DWL(t = \alpha^{OLS} + \beta^{OLS} z_j) - DWL(t = \bar{t})}{DWL(t = t_j^*) - DWL(t = \bar{t})} \\
&= 1 - \frac{DWL(t = \alpha^{OLS} + \beta^{OLS} z_j)}{DWL(t = \bar{t})} \\
&= 1 - \frac{-\frac{1}{2} \frac{\partial x_j}{\partial t_j} (1 - \frac{\partial mc_j}{\partial t_j}) SSR^{second-best}}{-\frac{1}{2} \frac{\partial x_j}{\partial t_j} (1 - \frac{\partial mc_j}{\partial t_j}) TSS^{second-best}} \\
&= R^2_{second-best}
\end{aligned}$$

where the equality follows from  $DWL(t = t_j^*) = 0$  and  $TSS^{second-best} = SSR^{uniform}$ . The coefficients of the regression  $t_j = \alpha^{OLS} + \beta^{OLS} z_j$  indicate the constrained-optimal tariffs.

### Welfare gain from a regression of *observed* marginal cost on tariff variables

We do not observe the benchmark marginal cost  $mc_j^*$ , but we do observe the marginal cost resulting from an unbiased *flat* tariff  $\bar{t}$ . However, the  $R^2$  of a regression of  $mc_j^*$  on the policy variables equals the  $R^2$  of a regression of the observed  $mc_j(\bar{t})$  on the policy variables, since  $mc_j(\bar{t})$  is an affine transformation of  $mc_j^*$ . To see this, denote the slope of marginal cost  $mc'_j$  and note that  $mc_j^* = t_j^*$ . We can write:

$$\begin{aligned}
mc_j(\bar{t}) &= mc_j^* + mc'_j(x_j(\bar{t}) - x_j(t_j^*)) \\
&= mc_j^* + mc'_j \frac{\partial x_j}{\partial t_j} (\bar{t} - t_j^*) \\
&= mc_j^* + mc'_j \frac{\partial x_j}{\partial t_j} \bar{t} - mc'_j \frac{\partial x_j}{\partial t_j} t_j^* \\
&= mc_j^* + mc'_j \frac{\partial x_j}{\partial t_j} \bar{t} - mc'_j \frac{\partial x_j}{\partial t_j} mc_j^* \\
&= (1 - mc'_j \frac{\partial x_j}{\partial t_j}) mc_j^* + mc'_j \frac{\partial x_j}{\partial t_j} \bar{t},
\end{aligned}$$

where  $mc'_j \frac{\partial x_j}{\partial t_j} \bar{t}$  and  $mc'_j \frac{\partial x_j}{\partial t_j}$  are constants, so  $mc_j(\bar{t})$  is an affine transformation of  $mc_j^*$ . Hence, the  $R^2$  from a regression of observed marginal costs on policy variables recovers the relative efficiency gain of the tariff structure under consideration (over a flat tariff). In order to compute the constrained-optimal policy and the welfare loss in dollars, the transformation above can be used to construct  $mc_j^*$  and run a regression of  $mc_j^*$  on the policy variables.

## Proof of Proposition 4

We first show that under Assumptions 1, 2, 4, 5 and 6, the second-best policy is  $\alpha = 0$  and  $\beta = \bar{r}$ . Consider an OLS policy  $t_{js} = \alpha + \beta f_j$ . 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 16, the deadweight loss from this OLS policy is:

$$\begin{aligned}
-2\text{DWL}(t_j = \alpha + \beta f_j) &= \sum_{s=1}^S \sum_{j=1}^J \sum_{k=1}^J e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^S \sum_{j=1}^J \sum_{k=1}^J ((r_s - \beta)f_j - \alpha)((r_s - \beta)f_k - \alpha) \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^S (r_s - \beta)^2 \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \alpha \sum_{s=1}^S (r_s - \beta) \sum_{j=1}^J \sum_{k=1}^J (f_j + f_k) \frac{\partial x_{js}}{\partial t_{ks}} \\
&\quad + \alpha^2 \sum_{j=1}^J \sum_{k=1}^J \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \times \sum_{s=1}^S (r_s - \beta)^2, \tag{A.17}
\end{aligned}$$

where the third equality follows from Assumption 5 (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 6 (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  $\alpha^2$ ) is zero. Second, under quasilinearity, the demand matrix is symmetric, so  $\sum_{j=1}^J \sum_{k=1}^J f_j \frac{\partial x_{js}}{\partial t_{ks}} = \sum_{j=1}^J \sum_{k=1}^J f_k \frac{\partial x_{js}}{\partial t_{ks}}$ , and under Assumption 6 (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  $\alpha$  are equal to zero, and deadweight loss is reduced to only the first term involving  $\beta$ .

The  $\alpha$  terms cancel because  $\alpha$  is just a lump sum transfer between the government and consumers. Given Assumption 6, with revenue-recycling, the constant has no effect on welfare and the optimal  $\alpha$  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.17) with respect to  $\beta$ . The first-order condition is:

$$\frac{\partial \text{DWL}}{\partial \beta} = \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} \times \sum_{s=1}^S (r_s - \beta) = 0. \tag{A.18}$$

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.17 for  $t_j = \bar{r} f_j$ :

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

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

$$\begin{aligned}
-2DWL(t = \bar{t}) &= \sum_{s=1}^S \sum_{j=1}^J \sum_{k=1}^J e_{js} e_{ks} \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^S \sum_{j=1}^J \sum_{k=1}^J (r_s f_j - \bar{r}\bar{f})(r_s f_k - \bar{r}\bar{f}) \frac{\partial x_{js}}{\partial t_{ks}} \\
&= \sum_{s=1}^S r_s^2 \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^J \sum_{k=1}^J \frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2 \bar{f} \times \sum_{j=1}^J \sum_{k=1}^J (f_j + f_k) \frac{\partial x_{js}}{\partial t_{ks}} \\
&= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}},
\end{aligned} \tag{A.20}$$

where, as detailed above in the derivation of Equation A.17, the third equality follows from common markets (Assumption 5) and the latter two terms are zero because of the no outside good assumption (Assumption 6).

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

$$\begin{aligned}
\frac{DWL(t = \bar{r}f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})} &= \frac{\frac{1}{2}S \times (E(r_s^2) - \text{var}(r_s)) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{\frac{1}{2}S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}} \\
&= 1 - \frac{\text{var}(r_s)}{E[r_s^2]}.
\end{aligned} \tag{A.21}$$

Moving to the fixed-effects intuition follows directly from de-meaning the regression at the level of the (region  $s$ ) fixed-effects and calculating  $R^2$  in the resulting regression (the definition of the within- $R^2$ ). This amounts to a regression of  $\phi_{js} - r_s \bar{f}$  on  $f_j - \bar{f}$ . The total sum of squares is:

$$\begin{aligned}
TSS &= \sum_{s=1}^S \sum_{j=1}^J (\phi_{js} - r_s \bar{f})^2 = \sum_{s=1}^S \sum_{j=1}^J (r_s (f_j - \bar{f}))^2 = \sum_{s=1}^S r_s^2 \sum_{j=1}^J (f_j - \bar{f})^2 \\
&= S \times E[r_s^2] \times \sum_{j=1}^J (f_j - \bar{f})^2.
\end{aligned} \tag{A.22}$$

A standard derivation of OLS shows that the slope is  $\bar{r}$ , and the constant is 0. The OLS residuals from the regression are therefore given by  $\phi_{js} - r_s \bar{f} - \bar{r}(f_j - \bar{f}) = r_s f_j - r_s \bar{f} - (\bar{r} f_j - \bar{r} \bar{f}) = (r_s - \bar{r})(f_j - \bar{f})$ . Now compute the sum of squared residuals:

$$\begin{aligned}
SSR &= \sum_{s=1}^S \sum_{j=1}^J [(r_s - \bar{r})(f_j - \bar{f})]^2 = \sum_{s=1}^S (r_s - \bar{r})^2 \sum_{j=1}^J (f_j - \bar{f})^2 \\
&= S \times \text{var}(r_s) \times \sum_{j=1}^J (f_j - \bar{f})^2.
\end{aligned} \tag{A.23}$$

Now we can compute within- $R^2$  as:

$$\text{within-}R^2 = 1 - \frac{SSR}{TSS} = 1 - \frac{S \times \text{var}(r_s) \times \sum_{j=1}^J (f_j - \bar{f})^2}{S \times E[r_s^2] \times \sum_{j=1}^J (f_j - \bar{f})^2} = 1 - \frac{\text{var}(r_s)}{E[r_s^2]}. \tag{A.24}$$

## Proof of Corollary 2

As demonstrated in Equation A.20, the deadweight loss for a constant unbiased tax  $t = \bar{t} = \bar{r}\bar{f}$  equals:

$$\begin{aligned}
-2\text{DWL}(t = \bar{t}) &= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^J \sum_{k=1}^J \frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2 \bar{f} \times \sum_{j=1}^J \sum_{k=1}^J (f_j + f_k) \frac{\partial x_{js}}{\partial t_{ks}} \\
&= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^J \sum_{k=1}^J \frac{\partial x_{js}}{\partial t_{ks}} - 2S\bar{r}^2 \bar{f} \times \sum_{j=1}^J \sum_{k=1}^J f_j \frac{\partial x_{js}}{\partial t_{ks}}.
\end{aligned} \tag{A.25}$$

where the second and third terms do not cancel if Assumption 6 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  $\theta_j$  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,  $\theta_j$  is the total market size effect for a change in tax rate  $t_j$ ; it is the ratio of the sum of cross effects (increases in quantity for other products that results from raising price  $j$ ) to the own 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.25](#):

$$\begin{aligned}
-2\text{DWL}(t = \bar{t}) &= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^J \left( \frac{\partial x_{js}}{\partial t_{js}} + \sum_{k \neq j} \frac{\partial x_{js}}{\partial t_{ks}} \right) \\
&\quad - 2S\bar{r}^2 \bar{f} \times \sum_{j=1}^J \sum_{k=1}^J f_j \left( \frac{\partial x_{js}}{\partial t_{js}} + \sum_{k \neq j} \frac{\partial x_{js}}{\partial t_{ks}} \right) \\
&= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + S\bar{r}^2 \bar{f}^2 \times \sum_{j=1}^J (1 - \theta_j) \frac{\partial x_{js}}{\partial t_{js}} \\
&\quad - 2S\bar{r}^2 \bar{f} \times \sum_{j=1}^J f_j (1 - \theta_j) \frac{\partial x_{js}}{\partial t_{js}} \\
&= S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - S\bar{r}^2 \bar{f}^2 \times J(1 - \bar{\theta}) \frac{\partial x_j}{\partial t_j}, \tag{A.26}
\end{aligned}$$

where the last equality follows from Assumption [3](#) and two other simplifying assumptions that we make here:  $f_j$  and  $\theta_j$  are uncorrelated, and  $\theta_j$  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 so the summation terms collapse to expressions of the mean.

In this case, the sign of the second term in Equation [A.26](#) 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  $\theta$ .) When  $\bar{\theta} = 1$ , the second term in Equation [A.26](#) 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 ( $\frac{\partial x_j}{\partial t_j}$  is more negative). Note that it would be unusual 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:

$$\begin{aligned}
& \frac{DWL(t = \bar{r}f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})} \\
= & \frac{-\frac{1}{2}S \times \text{var}(r_s) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} + \frac{1}{2}S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \frac{1}{2}JS\bar{r}^2\bar{f}^2(1 - \bar{\theta}) \frac{\partial x_j}{\partial t_j}}{\frac{1}{2}S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - \frac{1}{2}JS\bar{r}^2\bar{f}^2(1 - \bar{\theta}) \frac{\partial x_j}{\partial t_j}} \\
& = 1 - \frac{S \times \text{var}(r_s) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - JS\bar{r}^2\bar{f}^2(1 - \bar{\theta}) \frac{\partial x_j}{\partial t_j}}.
\end{aligned} \tag{A.27}$$

It follows directly from Equation A.27 that  $\bar{\theta} < 1$  implies that the  $R^2$  overstates the fraction of the welfare gain from the Pigouvian benchmark achieved by this second-best policy over a policy of a constant tax on all products (note that the numerator and first term of the denominator in the fraction below are negative):

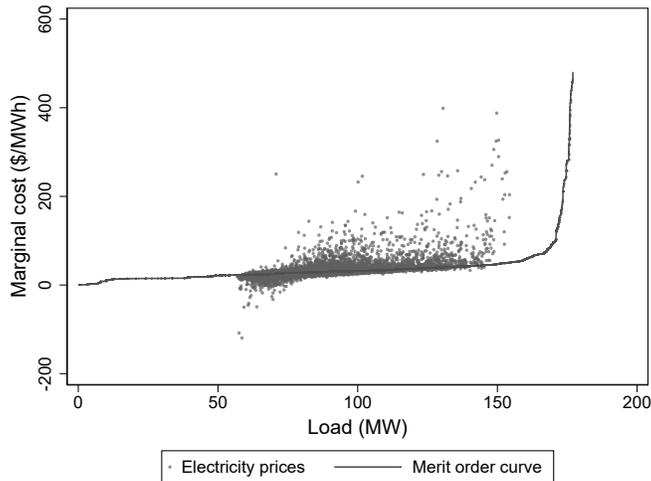
$$\begin{aligned}
\frac{DWL(t = \bar{r}f_j) - DWL(t = \bar{t})}{DWL(t = \phi) - DWL(t = \bar{t})} &= 1 - \frac{S \times \text{var}(r_s) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}} - JS\bar{r}^2\bar{f}^2(1 - \bar{\theta}) \frac{\partial x_j}{\partial t_j}} \\
&< 1 - \frac{S \times \text{var}(r_s) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}}{S \times E(r_s^2) \times \sum_{j=1}^J \sum_{k=1}^J f_j f_k \frac{\partial x_{js}}{\partial t_{ks}}} = 1 - \frac{\text{var}(r_s)}{E[r_s^2]} = \text{within-}R^2.
\end{aligned} \tag{A.28}$$

## B Impact of Cross-Price Substitution and Convexity in Application 2

In this Appendix we investigate the potential bias of  $R^2$  in the presence of cross-price derivatives and marginal cost convexity in the electricity pricing application. To simulate welfare, we use hourly data on wholesale electricity prices from 2012 in what is called the Pennsylvania-New Jersey-Maryland market (PJM). Data on real-time wholesale prices and load for PJM are available from their website. We use the reported “system price,” which measures a weighted average wholesale price for a given hour across all nodes within PJM, and “system load” data, the sum of demand within PJM in a given hour. For analyzing the robustness to marginal cost convexity we also use information on PJM’s 2012 merit order. These data were purchased from SNL Energy and report plant-level information such as fuel type, capacities, fuel costs, variable operating and maintenance costs, and emission allowance costs for all generating sources within PJM’s territory selling into the wholesale market. Figure B.1 shows a scatterplot of hourly wholesale prices and loads for 2012.<sup>43</sup>

<sup>43</sup>The reader will note the presence of negative electricity prices. These can arise for two reasons that interact with the fact that electricity cannot simply be disposed of. First, there is a federal production tax credit for wind of \$23/MWh. Therefore wind generators are willing to continue to produce as long as the price is above negative

**Figure B.1:** Electricity Prices and Merit Order in the PJM market at Varying Loads



## B.1 Cross-Price Substitution

Here we relax Assumption 3 in the context of electricity pricing, allowing a form of intertemporal substitution in demand that may violate our assumption that cross effects are proportional to own effects. When substitution is concentrated among neighboring hours, serial correlation in the pricing scheme will create violations of Assumption 3. For example, the sequential hours within a “peak” period will all be priced the same and are also stronger substitutes in demand. As described in Section 2.4, this pattern will push the welfare fraction evaluated using the full formula below the value of  $R^2$ . At the same time, serial correlation could also create correlation between cross-price derivatives and pricing errors (for example if a generator is offline for a few days), pushing the welfare fraction above  $R^2$ .

We explore a wide range of cross-price derivatives and find that, in spite of the theoretical potential for bias, relatively little bias appears in this application. Table B.1 reports results from a simulated demand system where cross-price elasticities are set to reproduce the shoulder period substitution estimated in Andersen et al. (2017). That paper estimates substitution to a period between 2 and 12 hours on either side of the price change and finds an average cross-price shift in demand equal to 29% of the size of the own-price demand response. We simulate this effect in our setting by inserting cross-price derivatives into the demand matrix used in Section 4, adjusting a matrix of share elasticities until the cross-price derivatives are 29%, symmetry is imposed, and substitution to the outside good is such that the rows (and therefore columns) sum to zero.<sup>44</sup> Note that because the cross-price derivatives require the log of prices, we are unable to account for

---

\$23/MWh. Second, traditional steam power plants face startup costs. This implies that during periods of low demand that are expected to be followed by high demand, generators are willing to pay consumers to take their electricity to avoid these startup costs in the future.

<sup>44</sup>We apply an iterative search process, alternating between scaling the cross-price derivatives and imposing the adding-up conditions, until the target cross-price share elasticities are reached.

negative prices. Thus, we impose a minimum price of \$1/MWh in our calculations. Because of this the  $R^2$  measures reported in Table B.1 are slightly different from those reported in the main text.

**Table B.1:** Substitution Across Hours Following Andersen et al. (2017)

Pricing scheme	Welfare formula including cross-price effects:					
	(1) $R^2$	(2) None	(3) 2 hours	(4) 3 hours	(5) 6 hours	(6) 12 hours
On vs off peak FE	0.041	0.041	0.033	0.031	0.033	0.045
HOD FE	0.135	0.135	0.126	0.127	0.134	0.148
HOD & DOW FE	0.153	0.153	0.141	0.142	0.148	0.162
HOD & MOY FE	0.193	0.193	0.177	0.177	0.181	0.192
HOD, MOY & DOW FE	0.211	0.211	0.192	0.191	0.194	0.205
HOD x MOY FE	0.297	0.297	0.282	0.283	0.291	0.306
HOD x MOY x DOW FE	0.422	0.422	0.411	0.412	0.418	0.429

Note: In columns (3) through (6) total substitution to the shoulder period is set equal to 29%, the mean estimate in Andersen et al. (2017). The shoulder period in their experiment varied between 2 and 12 hours depending on the treatment and so we present results spanning this range. HOD, DOW and MOY refer to hour of day, day of week, and month of year, respectively.

We find that the bias overall is quite small in magnitude and that the direction depends on the number of hours in the shoulder period. Even though the intertemporal substitution estimated in Andersen et al. (2017) is substantial (and much larger than the shift found in other pricing experiments), the overall bias in the  $R^2$  measure equals at most two percentage points.

**Table B.2:** Substitution Across Hours Using Estimates from Ata, Duran, and Islegen (2016) and Jessoe and Rapson (2014)

Pricing scheme	$R^2$	Cross-price effects considered:			
		None	(1)	(2)	(3)
On vs off peak FE	0.041	0.041	0.036	0.051	0.056
HOD FE	0.135	0.135	0.130	0.147	0.152
HOD & DOW FE	0.153	0.153	0.146	0.168	0.175
HOD & MOY FE	0.193	0.193	0.184	0.215	0.224
HOD, MOY & DOW FE	0.211	0.211	0.200	0.236	0.246
HOD x MOY FE	0.297	0.297	0.289	0.317	0.326
HOD x MOY x DOW FE	0.422	0.422	0.416	0.436	0.441

Notes: Column (1) sets cross-price effects equal to approximately 9% using estimates from the model in Ata, Duran, and Islegen (2016). Column (2) sets cross-price effects to negative 68%, the average of spillovers across treatments in Jessoe and Rapson (2014). Finally, column (3) uses the maximum spillover estimated for any treatment in Jessoe and Rapson (2014), slightly more than 100% of the own-price derivative in the treated hours. HOD, DOW and MOY refer to hour of day, day of week, and month of year, respectively.

Table B.2 explores two additional calibrations. First, we draw from estimates in Ata, Duran, and Islegen (2016) who measure substitution two hours on either side of a price change. They find more modest cross-price effects than those in Andersen et al. (2017), with substitution offsetting roughly 9% of reductions during a period of increased prices. The biases are correspondingly smaller. Finally, we consider the results in Jessoe and Rapson (2014), where they find that information

treatments combined with price changes can introduce substantial spillovers: cross-price derivatives are negative so that changes in an hour with a price change are amplified in the two hours before and after. This particular substitution pattern is large (greater than 100% in one of their treatments), but even so we find the bias this causes for the  $R^2$  is small relative to the variation across policies. We also note that the ranking of policies using  $R^2$  and the full model is identical in all simulations we have performed.

## B.2 Convexity Simulation for Application 2

Here we relax the linearity assumption in supply. We make use of engineering data on capacities and engineering marginal costs described above for each of the plants within the PJM service area, the so-called merit order. Figure B.1 superimposes the 2012 merit order on the scatterplot of wholesale prices and loads in 2012. There is a clear upward slope, but also variation in realized cost at any given level of load. The merit order provides information on the average availability and costs of power plants operating in the market, so in any one hour prices can be above or below this. The difference between the real-time price and the engineering estimate of marginal cost can be due to variation in production from renewables, periods when other generators are offline due to maintenance, renewable energy subsidies or transmission congestion that limits the amount of electricity a low-cost generator can sell to particular locations.

Using data on the merit order, we simulate deadweight loss from a particular intermediate pricing scheme in each hour. We compare the welfare implications of seven pricing schemes in Table 1. To generate each intermediate set of prices we regress the hourly wholesale prices on the set of dummy variables defined by the intermediate pricing scheme. The fitted values from these regressions represent prices in each scheme.<sup>45</sup>

We also require both supply and demand to simulate deadweight loss from each of these intermediate pricing schemes in a given hour. Consistent with our theoretical results, we assume demand is (locally) linear. We calibrate demand such that the average elasticity across the year is 0.1. We use two supply curves. First, we estimate a quadratic supply curve through the merit order via least squares.<sup>46</sup> Second, we use the actual step-function merit order.

In each hour, we project our linear demand curve from the real-time wholesale price. Because the hourly prices do not lie on the estimated supply curves for the reasons discussed above (e.g., power plant outages), we horizontally shift the quadratic and step-function supply curves so that they intersect with the hourly wholesale price. This implicitly assumes that the supply curve has shifted due to changes in production from infra-marginal resources.<sup>47</sup> Figure B.2 illustrates this calculation (after shifting the supply curve) for a sub-optimal price that is lower than the real-time

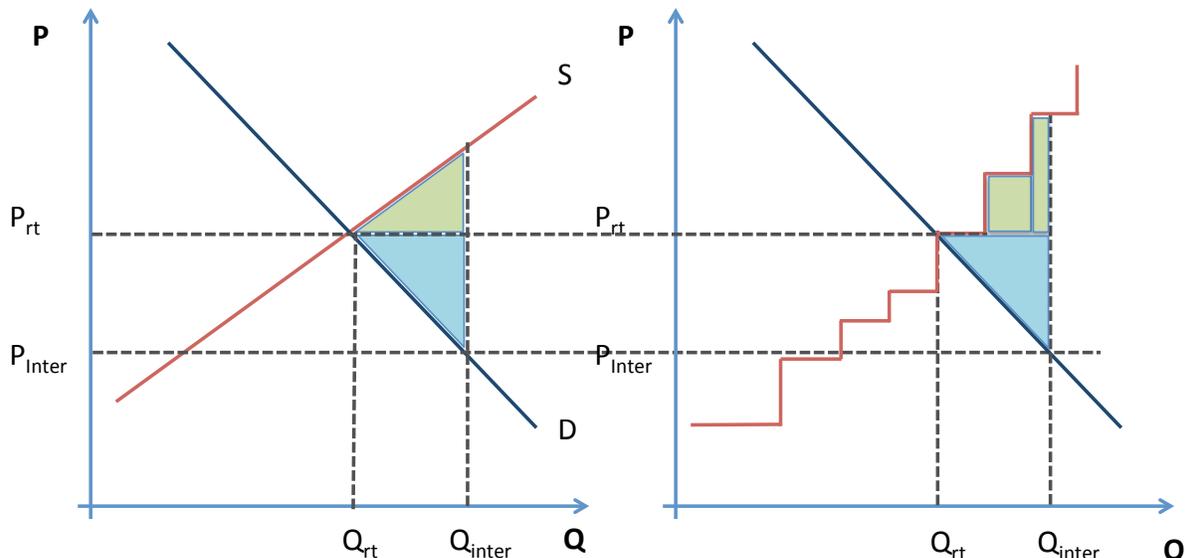
<sup>45</sup>We do not weight the regression by load, but the results are qualitatively similar when we do.

<sup>46</sup>We include only the squared term and a constant in the regression. Including the first-order term leads to the unrealistic result of marginal cost initially declining.

<sup>47</sup>An alternative method would be to vertically shift the supply curves. Here, the implicit assumption would be that the marginal cost of the marginal plant and all other plants increase by the same amount. One advantage of horizontally shifting the supply curves is that it uses only data on wholesale *prices*, not needing information on wholesale loads.

price for linear demand and linear and step-function supply. Given linear demand, the deadweight loss coming from consumers will always be a triangle with a height equal to the price difference and a width defined by the slope of demand. For the quadratic supply, the deadweight loss arising from the supply side integrates from the baseline quantity over the change in quantity defined by the price differences and the slope of the demand curve. For the step-function supply, deadweight loss coming from the supply side is the sum of the series of rectangles over this change in quantity.<sup>48</sup>

**Figure B.2:** Deadweight Loss Calculations under Intermediate Pricing Schemes



Note:  $P_{inter}$  is the second-best, or intermediate, price, against with  $P_{rt}$ , the real-time price, is evaluated.

For the vast majority of the hours (8,544 of the 8,784 hours),<sup>49</sup> the steps discussed above are sufficient for calculating deadweight loss. However, there are two cases where calculating deadweight loss using the step function in this way is not possible. The first is in hours where the observed real-time price is below the initial starting point of the merit order (\$0.55/MWh). The second is when the counterfactual demand under the intermediate pricing scheme ( $Q_{inter}$ ) falls beyond the vertical portion of the supply curve.

We considered a number of ways to deal with these “outlier” hours. For example, we could vertically shift the supply curve during these low price hours, or we could stretch out supply or assume some marginal cost for imports in hours where demand under the intermediate price is beyond the capacity of the system. Given the relatively small number of hours for which this is an issue, in the tables below we omit these 247 hours from both the deadweight loss calculations and the formation of the intermediate prices.<sup>50</sup>

<sup>48</sup>Under the step-function supply the real-time price can either be between steps or at the horizontal portion of the step. When the real-time price is at a horizontal section, we shift the step-function supply curve so that the observed price is at the beginning of the step as shown.

<sup>49</sup>Recall that 2012 was a leap year.

<sup>50</sup>To do this, we first use all of the hours to calculate intermediate prices and then identify the problematic hours.

Table B.3 reports the results from this procedure using the slightly smaller sample, as well as the  $R^2$  results for the full sample reported in Table 1 reported in the main text. Expectedly, because we are omitting price extremes, the share of deadweight loss eliminated is larger than in results reported in the main text. We find almost no bias using the quadratic supply curve for each of the seven pricing regimes. The bias under the step function supply is at most 10.4%, except for the first case where the absolute discrepancy is very small. While we have not shown this theoretically, for each of the pricing schemes  $R^2$  is slightly biased upward. In addition, the bias is fairly stable, in percentage terms, across the pricing schemes (except for the simple peak/off-peak tariff) ranging from 6.6% to 10.4%.

**Table B.3:** Comparison of  $R^2$  with Simulated Welfare Measures from Electricity Tariff Application

Pricing Regime	$R^2$ , complete	$R^2$ , restricted	Simulated, quadratic	Simulated, step-function
On vs. off peak FE	0.040	0.042	0.042	0.063
HOD FE	0.135	0.253	0.257	0.280
HOD & DOW FE	0.153	0.278	0.282	0.310
HOD & MOY FE	0.193	0.329	0.333	0.355
HOD, DOW & MOY FE	0.211	0.354	0.358	0.383
DOW x MOY FE	0.297	0.449	0.452	0.499
HOD x MOY x DOW FE	0.428	0.599	0.599	0.642
Number of observations	8,784	8,544	8,544	8,544

*Notes:* Data are from hourly wholesale electricity prices in the PJM market for 2012. Peak hours are defined as 2-6 p.m. Complete Sample uses all 8,784 hourly prices, while the restricted sample omits hours where price is below the minimum marginal cost and hours where the counterfactual demand under constant pricing exceeds the capacity available in the merit order. HOD, DOW and MOY refer to hour of day, day of week, and month of year, respectively.

We reiterate that our goal here is to verify the validity of our sufficient statistics approach; it is not to provide a full analysis of the welfare gains from particular pricing rules. A variety of caveats to our analysis exist. First, we are analyzing only one year. The welfare implications may vary across years for a number of reasons. Second, we are looking only at the welfare implications of prices constructed *within sample*. In practice, policy must set prices before real-time prices are realized. Third, this is only for one geographic market.

These caveats, however, increase the importance of being able to calculate the welfare implications of different pricing regimes across multiple markets and time periods with minimal modeling effort. For example, our method could be used to quickly estimate the optimal moving window of data that policymakers should use to set prices in advance. Or, policymakers could use our method to verify the external validity of welfare effects from one pricing rule in another market. Outside of our framework, such comparisons would have imposed a much higher modeling and data burden.

Next, we re-estimate the intermediate prices omitting these problematic hours. We have found that these second set of intermediate prices do not generate any additional problematic hours. We have also calculated the portions of deadweight loss defined for all hours (e.g., all of the deadweight loss from consumers and portions of the deadweight loss from supply) and the results are qualitatively similar.

## C Appendix for Application 3

### Data details

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. 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 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.<sup>51</sup>

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 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.<sup>52</sup> 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 U.S. for two years.

---

<sup>51</sup>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.

<sup>52</sup>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.

## Robustness: outliers, sampling and censoring

### Outliers

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 our main results, we have dropped observations for which VMT-at-death exceeds one million miles. To check whether our  $R^2$  results are sensitive to different treatments of observations with very high VMT-at-death, Table C.4 reports regressions that include all observations as well as regressions in which we winsorize the underlying micro-data at different VMT thresholds.<sup>53</sup> 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. The first two rows indicate that dropping observations with VMT above one million miles hardly affects  $R^2$ . Rows 3-6 indicate that, starting from the full sample, winsorizing at progressively lower VMT levels only slightly increases  $R^2$ . 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.<sup>54</sup> The WLS  $R^2$  rises from a baseline of 0.22 to a maximum of 0.30 for the same restriction. Our qualitative conclusions are therefore robust to outliers.

**Table C.4:** 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.

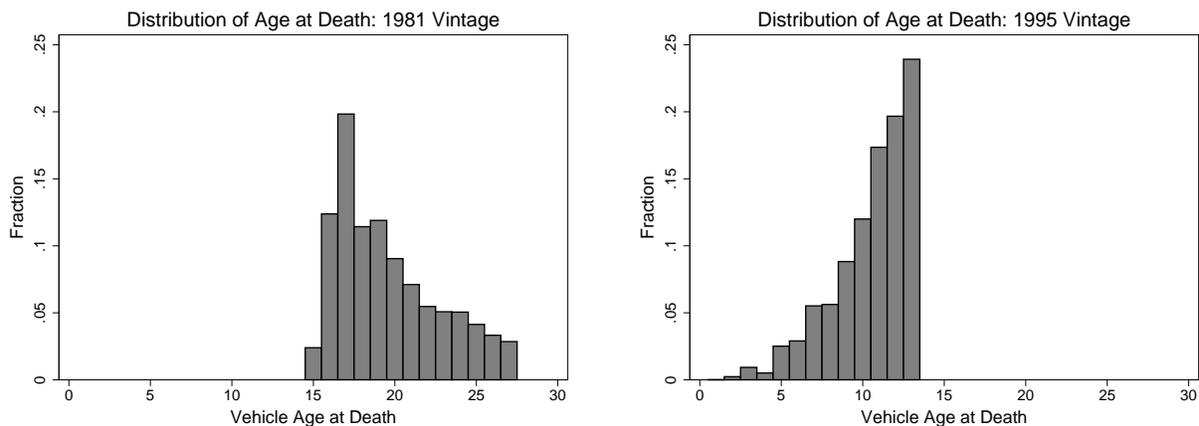
### Sampling variation

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

<sup>53</sup>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.

<sup>54</sup>Our data have a median  $VMT$  at scrappage of 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 lifetime mileage than would be true in other climate zones.

**Figure C.3:** The Distribution of Vehicle Age at Death for Different Vintages



retirements in our sample. We then bootstrap that sample and estimate the  $R^2$  many times. The mean estimate is 0.283, which is close to the 0.29 from the corresponding specification in Table 2. 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 0.06.

## Censorship

Censorship bias is a 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 C.3 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.<sup>55</sup> This censoring can create (non-classical) mismeasurement, which can be particularly problematic when comparing across cohorts. However, we will now show 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).

In Section 5.5 we provide a first indication that the bias from censoring may be limited. Here

<sup>55</sup>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.

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 national registration count data from RL Polk at the VIN10-prefix level. We use these uncensored data to compute annual scrap rates for each VIN10-prefix over the sample period 1999-2009 and then to fill in missing scrap rates in our main data wherever possible. 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).<sup>56</sup> Having extrapolated missing scrap rates (and, indirectly, missing vehicle retirements), we then impute missing VMT-at-death using a similar procedure. We first replace missing VMT-at-death for each age using VMT averages across VIN8-prefixes. For ages that are never observed at the VIN8-prefix level, we use the 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 C.5 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 retirements (including imputed ones) (panel 1). While this range is clearly above 0.22-0.29 (as reported in panel 2 of Table 2), the  $R^2$ s are still low from an absolute perspective. Panel 2 shows that when we restrict the sample to VIN10-prefixes for which we impute VMT-at-death for at most 12 ages, the range goes down to 0.23-0.34. This provides further evidence that censoring is unlikely to cause a large bias.

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.

---

<sup>56</sup>Specifically, we fit a fifth-order polynomial to the scrap rate by age pattern, and use this for imputing missing data.

**Table C.5:** Regression  $R^2$  Using Imputed Data

VIN10-prefix averages, model years 1988-1992	OLS	WLS
<b>Panel 1: VMT imputed for all models</b>		
All models	.44	.43
Models with $N \geq 200$	.45	.38
Models with $N_{imputed} \geq 400$	.47	.40
<b>Panel 2: Only models for which VMT is imputed for <math>\leq 12</math> ages</b>		
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  $VMT$  above one million miles are dropped. WLS uses the actual number of observed retirements  $N$  when the sample is selected based on  $N \geq 200$ . WLS uses the number of retirements including imputed retirements  $N_{imputed}$  when the sample is selected based on  $N_{imputed} \geq 400$ .

Specifically, we restrict the sample to models with  $N \geq 200$  and model years 1988-1992 and show how  $R^2$  changes as we progressively remove vehicles of older ages from the sample. Table C.6 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.<sup>57</sup> 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  $R^2$ .

### Policies with a bias

Our theoretical results assume that the policy-maker chooses an unbiased second-best tax scheme. When the policy is in fact biased, there will be, by definition, additional welfare losses. 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

<sup>57</sup> $R^2$  decreases from 0.37 for 10-year-old vehicles to 0.13 for 20-year-old vehicles.

**Table C.6:** 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  $VMT$  above one million miles are dropped. WLS weights the regressions by the number of observed retirements  $N$ .

the average tax wrong (“mean bias”).<sup>58</sup> In the context of our fuel-economy application, a downward bias in the mean tax rate would fail to shrink the car market by the optimal amount (market size becomes relevant in settings with an outside good). Revenue-neutral fuel-economy standards by definition set an average tax rate of zero across all cars; since the average externality is positive this constitutes a downward bias.<sup>59</sup> This creates an additional inefficiency at the extensive margin, such that  $R^2$  in the tables above can be interpreted as an upper bound on the relative welfare gain. 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. We show here the difference between this “naïve” linear tax from the actual second-best. We illustrate this in Figure C.4, which replicates Figure 4, but adds a line that represents the relationship between fuel consumption ratings and lifetime fuel consumption, if all cars (or trucks) were driven the same number of miles, which we set equal to the observed

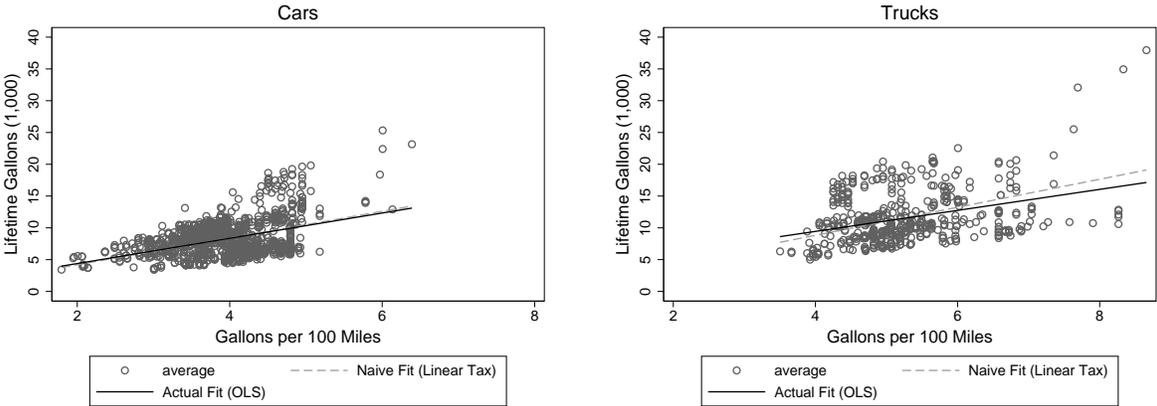
<sup>58</sup>In terms of the theory, a useful decomposition is to separate the mean bias in tax rates from their variance, which can be seen by rewriting deadweight loss in Equation 10 under Assumptions 1 to 3:

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

This illustrates that when there is bias in the tax rates, and variance in their errors, the 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 this equation, but (non-OLS) policies with a slope bias have a larger variance term.

<sup>59</sup>Graphically, this can be represented by shifting downward the linear tax schedule in Figure C.4. Because a revenue-neutral 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.

**Figure C.4:** 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* above one million 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.

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 C.4 shows that the naïve linear tax differs noticeably from the best linear tax for trucks, but that the difference for cars is small. This mis-pricing represents another source of inefficiency from ignoring heterogeneity in durability. In our case, this inefficiency turns out to be small, so we do not emphasize its implications, though it could be important in other contexts.