Pigou Creates Losers: On the Implausibility of Achieving Pareto Improvements from Efficiency-Enhancing Policies

James M. Sallee*
University of California, Berkeley and the NBER

May 6, 2019

Abstract

Economic theory predicts that efficiency-enhancing policy changes can be made to benefit everyone through the use of lump-sum transfers that compensate anyone initially harmed by the change. Precise targeting of compensating transfers, however, may not be possible when agents are heterogeneous and the planner faces constraints on the design of transfers. In this paper, I derive a necessary condition for an efficiency-enhancing policy to create a Pareto improvement that can be tested directly with data. The condition relates the size of efficiency gains to the degree of predictability between initial burdens and variables used to determine a transfer scheme. The main empirical application is to a gasoline tax to correct carbon emissions, with related results for other sin taxes also presented. Results indicate that it is infeasible to create a Pareto improvement from the taxation of these goods, and moreover that plausible policies are likely to leave a large fraction of households as net losers. The paper argues that the existence of these losers is relevant to policy design and may help explain the political challenges faced by many efficient policies. The paper concludes with several extensions related to this political economy motivation.

Keywords: Corrective taxation, externalities, equity
JEL: H23, Q58, L51

*The author thanks Karl Dunkle-Warner and Catherine Wright for excellent research assistance, Shanjun Li and Gilbert Metcalf for conference discussion, and Lint Barrage, Severin Borenstein, Lucas Davis, Don Fullerton, Kelsey Jack, Gary Libecap, Ethan Ligon, Dmitry Taubinsky and seminar participants at ASSA, Berkeley, the NBER and HERE conferences for helpful comments. The Sloan Foundation and the Hellman Fund at the University of California, Berkeley provided generous support. © 2017 James M. Sallee. All rights reserved.

1 Introduction

Why do efficient policies so often fail to gain political traction? Many policies are widely viewed as desirable by economists but unpopular with the public and unsuccessful in the policy process. Examples range from the pricing of pollution to repeal of the mortgage interest tax deduction to free trade. Several factors may lead to unpopularity of such policies, one of which regards the distribution of burdens they induce. Distributional concerns come in two varieties. In one, a policy is disliked because it is regressive and disproportionately affects low-income households. In the other, a policy imposes a substantial burden on a particular firm, set of firms, or group of voters who mobilize to block the policy.

In either case, economic theory provides a potential reply, which is that any such losers can be compensated. Any efficiency-enhancing policy, by definition, creates enough new surplus to compensate all losers. That is, any Kaldor-Hicks efficiency gain can be made into a Pareto improvement, if the right transfers are made in the background. A regressive tax could be combined with tax reform so as to preserve the desired income distribution, or any firms facing lost profits can be made whole.

The task of designing and implementing the right background transfers or reforms, however, may often be impossible when a planner is constrained to design a transfer function that is based on only some of the factors that determine initial policy burdens or their proxies. Constraints on the design of transfers could be due to information (e.g., preference heterogeneity is unobserved), demands for parsimony, administrative feasibility, or other factors.

This paper asks under what conditions is it possible to fully compensate losers from efficiency-enhancing policies via feasible lump-sum transfers. That is, when can efficiency-enhancing policies truly make everyone better off? In terms of theory, I derive a necessary condition that must be met for a Pareto improvement to be possible. This condition can be taken directly to data. Empirically, I test this condition for the case of externality-correcting taxes in the US, with a focus on motor fuels, when transfers are based on household demographics, income and geography. I find that the necessary condition fails and that a substantial fraction of households will be made net losers from externality-correcting policies. In brief, Pigouvian taxes create losers.

The basic idea is best illustrated via example. Consider a tax that increases efficiency by correcting an unpriced externality in the tradition of Pigou (1932). This policy creates a heterogeneous initial distribution of burdens across individuals depending on their taste for the taxed good. In most settings, the planner has enough revenue collected from the tax in order to compensate everyone for their loss through lump-sum transfers. But, compensating

everyone will require giving back the transfers in a targeted way. Targeting directly on consumption of the good itself will undo the desired corrective incentives, so the transfer must be based on factors like demographics, geography or income that are relatively inelastic and feasible for the planner. If the transfer function is not rich enough to precisely target transfers, then the planner will run out of available funds before fully compensating everyone. In this sense, the failure to create a Pareto improvement is due to a *prediction* problem; lump-sum transfers can only undo the distribution of burdens if they can be targeted precisely.

Summary of the paper: The paper proceeds by first laying out a theoretical framework that derives a necessary condition for it to be possible to turn an efficiency enhancing policy into a Pareto-improvement through transfers. I show that, for a Pareto improvement to be feasible, the variables that are used to determine the transfer scheme must precisely predict initial policy burdens, with the degree of precision related in a very simple way to the size of the average surplus gain created from the efficiency-enhancing action.

This condition can be directly tested with data, with the data requirements dependent on the policy in question. For a marginal increase in an externality-correcting tax, the initial burdens are measured directly by baseline consumption of the good, and the average welfare gain depends on an estimate of marginal external damages and a demand derivative. Thus, to check the condition for an externality-correcting tax, one needs (1) an estimate of the distribution of baseline consumption of the good, (2) knowledge of the correlation between baseline consumption and covariates that can be used in a transfer scheme, (3) an estimate of the own-price derivative, and (4) an estimate of the size of the externality.

To take the theory to data, I use the Consumer Expenditure Survey (CEX) to estimate the distribution of consumption of externality-creating goods and the correlation between consumption and covariates that could be used in transfer schemes. I combine this with estimates from the literature of the size of externalities and price derivatives. I initially focus on a gasoline tax used to correct carbon-related externalities. There is wide dispersion in consumption of gasoline across households, and only a modest fraction of this variation is correlated with variables that are likely to influence a transfer scheme, namely household structure, geographic location and income. Only about one-third of intrahousehold variation in annual gasoline expenditures is predictable by those variables, based on OLS and lasso models. Using conventional estimates of the externality gain achieved by a carbon tax, I conclude that the transfer scheme is nowhere close to precise enough to create a Pareto improvement (i.e., the necessary condition does not hold). Instead, in the most saturated model, I find that more than one-third of households are still net losers.

Additional variation can be explained with correlated endogenous variables. Specifically, vehicle ownership variables predict some of the remaining variation in gasoline expenditures,

as one would expect. Conditioning transfers to recycle the gasoline tax on vehicle ownership is clearly problematic in terms of incentives, but even using these variables only pushes the explanatory power up by a modest amount and leaves a large number of losers. (The preferred schemes involving demographics and income create incentive effects, as those characteristics will, to varying degrees, respond to the transfer scheme. In abstracting from these distortions, I am painting an optimistic picture for targeting, which still falls far short of creating a Pareto improvement. Below I discuss a second-best scheme that takes these distortions into account.)

I then show that the degree of predictability is no better for other externality-causing goods measured in the CEX, namely natural gas, electricity, alcohol and tobacco. I interpret this as evidence that it will be infeasible to create a Pareto improvement from corrective taxes on these goods, even when a planner uses an implausible amount of information to create an unrealistically flexible lump-sum transfer scheme.

Contributions and relationship to the literature: The theory and prediction exercise in this paper are focused on the narrow question of whether it is possible to compensate all losers from a corrective tax. But I also aim to make a broader point about how an empirical prediction problem lies at the heart of traditions in public finance that suggest efficiency and distributional concerns can be separated in policy analysis. The final section of the paper offers several related extensions.

A tradition in economics going back at least to Musgrave (1959) suggests that efficiency and equity concerns can often be conceptually divided. Given tools that can tilt the balance between rich and poor, like a progressive income tax, a policymaker should ensure market efficiency, and then simply dial up (or down) the levers that determine the income distribution to achieve the desired resource allocation in society. This is an extremely useful modeling device, and it is favored by many who study second-best tax design (e.g., Kaplow 2004). A literature in public finance explores the separability of efficiency-enhancing policies, including Pigouvian taxes, in second-best constrained environments (e.g., Gauthier and Laroque 2009; Kaplow 2012). This theoretical literature has noted that preference homogeneity is a critical assumption in their models, but little empirical work follows up by asking how these ideas can be implemented when there is some heterogeneity. Closely related is a seminal result of optimal tax theory that the distributional implications of a commodity tax are irrelevant in the presence of a nonlinear income tax (Atkinson and Stiglitz 1976). This likewise requires preference homogeneity (Saez 2002). This paper comments on these theoretical traditions by (1) deriving theoretical conditions that demonstrate when Pareto improvements are possible in the presence of some preference heterogeneity, (2) empirically testing the degree to which transfers can be adequately targeted so as to undo the initial distributional burdens of a

class of policies, and (3) demonstrating the relationship between heterogeneity and empirical prediction in achieving separation.

This paper bears an apparent relationship to several strains of literature in the theory of taxation, but it ultimately deals with different concerns. First, in being concerned with the correlation of tax burdens with covariates, the paper is related to the literature on tagging and targeting that follows Akerlof (1978), which considers how observable characteristics can be used to reduce distortionary tax incentives. Second, in being driven by a root information problem, this paper bears some relation to the literature begun by Mirrlees (1971), in which, if the planner could directly observe everyone's ability level, the optimal tax system would be nondistortionary. In my setting, if the planner could directly observe preference heterogeneity (and all other primitives that determine consumption of the externality-causing good), then Pareto improvements will be straightforward. Third, in considering optimal tax and transfer schemes to correct externalities, this paper is related to a literature—starting with the seminal work of Sandmo (1975) and with key contributions including Bovenberg and van der Ploeg (1994); Cremer, Gahvari, and Ladoux (1998, 2003); Jacobs and de Mooij (2015)—that derives second-best taxes on externality-creating commodities in order to maximize social welfare.

All three of these literatures are focused on how to derive second-best policies that minimize distortions caused by tax and transfer systems. My objective is different, at least proximately. My goal is to characterize the ways that imperfect information, which results in imperfect targeting/tagging, limits the planner's control over the final distribution of outcomes induced by an efficiency-enhancing reform. My empirical exercise is closer to the literature on targeting on observables prominent in the development literature, where the goal is to use readily measured proxies for wealth to target social programs. (See Coady, Grosh, and Hoddinott (2004) for a review.) The question at hand in designing the lump-sum transfer schemes is not maximization of social welfare (though that is the deeper reason why efficiency-enhancing policies are undertaken to begin with), but rather how to compensate the losers from the efficient scheme, with an eye on political economy, as explained next.

Should we be concerned about creating a Pareto improvement, or is it a red herring? Pareto efficiency vis-à-vis the status quo is quite distinct from social welfare maximization. If one begins with the objective of maximizing social welfare, there is no reason to prioritize the status quo resource allocation in society, so fussing over Pareto improvements is largely a distraction. The motivation for seeking Pareto improvements in this paper is instead a practical one. The political process tends to favor the status quo over changes, and as such, effecting change requires satisfying a great many people. That is, a utilitarian planner would gladly accept a policy that benefits most people, but causes modest harm to the remainder.

But, in practical terms, even small numbers of losers can create substantial political obstacles, consistent with the logic of collective action (Olson 1965, 1982). Empirically, this paper suggests that even implausibly well-designed schemes will leave large fractions of households as net losers. With this in mind, the final section of this paper suggests several ways that an empirical prediction problem can be adapted so as to inform the design of a "politically optimized" transfer scheme that accompanies a Pigouvian tax. This perspective focuses on the popularity of a policy based on its final distribution of impacts. Of course, popular policies need not be adopted, and a vast literature in political science and political economy studies various reasons why elected officials may behave differently from the desire of voters. I make no attempt to trace out all of this here. My claim is simply that the final distribution of impacts of a policy among individuals is a relevant, but not the only, factor in the policy-making process.

The paper's core negative result about Pigouvian taxes begs the question of whether a conceptually different approach to solving the problem of externalities would have better characteristics in terms of Pareto improvements relative to the status quo. If externalities are resolved via private negotiation as suggested by Coase (1960), then all parties must be made better off. But, the same information problems at issue here will often confound Coasian solutions. With asymmetric information about heterogeneity, private bargaining is subject to well known inefficiencies; some surplus improving trades will not take place. For non-excludable externalities affecting large numbers of actors as considered here, Coase's solution is also subject to free-riding problems. Overcoming free-riding requires the design of a collaboration mechanism, which will be hampered by private information about heterogeneity. Solutions may be possible in some situations (Ostrom 1990), but transaction costs likely limit the efficacy of this approach for large environmental problems like greenhouse gases. Below, I also argue that traditional mechanism design solutions, like Vickrey-Groves-Clarke mechanisms or implementation schemes following Varian (1994), do not solve the Pareto improvement conundrum.

In terms of the empirical application, this paper contributes to an existing literature on the distributional impacts of gasoline taxes (e.g., Poterba 1991; West 2004) and carbon taxes (e.g., Hassett, Mathur, and Metcalf 2009; Grainger and Kolstad 2010; Dinan 2012; Mathur and Morris 2014; Metcalf 2009; Burtraw, Sweeney, and Walls 2008; Williams, Gordon, Burtraw, Carbone, and Morgenstern 2015). That work has been overwhelmingly focused on measuring *average progressivity/regressivity* of taxes, whereas this paper is sharply focused on heterogeneity in policy burdens conditional on income and the degree to which that heterogeneity can be controlled via a transfer scheme.

A smaller recent literature does quantify heterogeneity in policy burdens conditional on

income. Rausch, Metcalf, and Reilly (2011) use the Consumer Expenditure Survey (CEX) to characterize the overall progressivity of carbon pricing, accounting for both consumption and income channels. Pizer and Sexton (2019) analyze the CEX and similar data from the United Kingdom and Mexico to show box plots that depict the range of energy consumption within income deciles. Fischer and Pizer (2019) explore how attention to horizontal equity influences a comparison between energy-pricing schemes and a performance standard. Cronin, Fullerton, and Sexton (2019) link the CEX to income tax data to explore a variety of revenue recycling mechanisms and quantify the variation in burdens that remains, taking into account fine-grained differences in income sources. Davis and Knittel (2019) show the heterogeneity in policy impacts of fuel-economy standards across different households in the same income decile in what is otherwise a study of average progressivity.[1]

These papers provide several initial results that are important for the development of a full analysis of heterogeneity in the incidence of energy policies. All demonstrate that there is significant heterogeneity in baseline energy consumption within households that have similar income, which are consistent with the descriptive facts I document here. Only Cronin, Fullerton, and Sexton (2019) link their study of heterogeneity to revenue redistribution schemes. They model several realistic schemes for revenue redistribution using detailed administrative tax records to show how the distribution of burdens depends on the use of revenue. I complement their approach first by modeling alternative transfer schemes that are explicitly designed to reduce heterogeneity in burdens, and second by providing a theoretical framework that demonstrates under what conditions revenue redistribution could plausibly achieve a Pareto improvement.

Many prior studies have discussed compensation schemes from externality-correcting taxes, and careful writers do sometimes note that schemes that achieve average redistributional goals will nevertheless create some losers (e.g., Metcalf 2018, p.98). Cronin, Fullerton, and Sexton (2019) and Fischer and Pizer (2019) both conjecture that, when there is a great deal of heterogeneity in baseline energy usage, it will be impossible to design transfer schemes to make everyone better off. My model offers a way to confirm their conjectures, and to show how much heterogeneity can exist before true Pareto transfers become infeasible.[2] I turn

[1] Some of this literature invokes the concept of horizontal equity. This concept has come under criticism as a normative criterion (Kaplow 1989). Here, I am concerned with horizontal equity, but not for normative reasons. This paper's argument is that horizontal equity matters for preventing the creation of losers and thereby improving political acceptability.

[2] Another related strand of literature focuses on compensating producers who are harmed by environmental regulation (Bovenberg and Goulder 2001; Bovenberg, Goulder, and Gurney 2005; Goulder, Hafstead, and Dworsky 2010). Most of that literature focuses on average impacts by sector or consumer group and does not delve into the heterogeneity that is the core of this study, though Burtraw and Palmer (2008) do consider individual power plants in an examination of the impacts on the electricity sector.

now to a description of that model.

2 A model of Pareto transfers

I begin with a treatment of the problem of achieving a Pareto improvement from a generic efficiency-enhancing policy. I then interpret the model for the case of an externality-correcting tax. My goal is to derive a necessary condition that must be met for a Pareto improvement from efficiency-enhancing policies to be possible that can be taken directly to data.

Costs, benefits and revenue: Consider some policy action that will create heterogeneous burdens, produce efficiency gains, and raise some revenue. Let heterogeneous agents indexed $i = 1, ..., N$ be the ones who bear initial burdens from the policy. Burdens are denoted c_i, with $\sum_i c_i = C$ so that $C > 0$ is the total initial cost of the policy. The set of agents who bear direct policy burdens are referred to as *in the market*.

The action yields efficiency gains of value $G > 0$. The gains enjoyed by agents who bear the policy burden are denoted g_i. Some fraction η of total gains go to the agents who bear the burden of the policy so that $\eta G = \sum_i g_i$. Gains are assumed to be weakly positive ($g_i \geq 0$). This assumption is not intended to be economically substantive, but it is used in the algebraic proofs below. It is convenient to characterize the welfare gains per agent, so I use $\bar{g} = \eta G/N$ to denote the average welfare gain enjoyed by agents in the market.[3]

The policy raises some revenue, denoted $R > 0$. Revenue can be redistributed through a transfer scheme based on exogenous covariates, $\mathbf{X}_i$. The transfer scheme is denoted $T(\mathbf{X}_i)$. The budget constraint requires that total transfers given out to agents is no greater than revenue $\sum_i T(\mathbf{X}_i) \leq R$. Note that $T(\mathbf{X}_i)$ can be negative—that is, the transfer can be a tax for some individuals. As discussed below, it is straightforward to accommodate the use of additional revenue that might be available to fund transfers.

The average *funding gap* is the per person difference between revenue raised and cost, denoted $\bar{\Delta} \equiv (C - R)/N$. This gap can be positive, negative or zero. A positive gap implies that the policy imposes costs that exceed revenue.

A Pareto improvement: For an efficiency-enhancing policy, total costs must be less than total benefits plus revenue: $C < R + G$. A Pareto improvement occurs when the analogous condition holds at the individual level for each individual, not just on average.

[3]Below, it is assumed the gains accruing outside the market—which are equal to $(1 - \eta)G$—cannot be taxed and used to compensate those in the market. Practically, this restriction can be relaxed by simply scaling up η: i.e, assuming that $\eta = 1$ is tantamount to assuming all gains accruing outside the market can be captured by the planner.

Thus, including the budget constraint, a Pareto improvement occurs when:

$$c_i < T(\mathbf{X}_i) + g_i \ \forall i \quad \text{and} \ \sum_i T(\mathbf{X}_i) \leq R.$$

To achieve a Pareto improvement, one must design a transfer scheme that delivers bigger transfers to those with bigger initial burdens. Intuitively, the goal will be to target transfers to offset burdens. Accordingly, I refer to the gap between initial burdens and the transfer $c_i - T(\mathbf{X_i})$ as the *targeting error*. A Pareto improvement requires that the targeting error be smaller than the efficiency gains ($c_i - T(\mathbf{X}_i) < g_i$) for all agents.

The main result: Condition 1 is a necessary condition for a Pareto improvement to be possible. That is, when the condition fails, a Pareto improvement is impossible. The condition directly motivates empirical analysis.

Condition 1. *Let c_i be the private burdens from a policy, N be the number of agents in the market, $T(\mathbf{X}_i)$ be a transfer scheme, $\bar{\Delta}$ the average funding gap, and $\bar{g}$ be the average efficiency gain accruing to those in the market. A Pareto improvement is not possible if the average absolute targeting errors exceed twice the average efficiency gain in the market minus the average funding gap; i.e., if*

$$\frac{1}{N} \sum_i |c_i - T(\mathbf{X}_i)| > 2\bar{g} - \bar{\Delta}$$

then a Pareto improvement is not possible.

Intuitively, condition 1 illustrates the relationship between the size of the surplus gain and the ability of a policy to precisely target transfers based on initial burdens. The left-hand size of the inequality is the average size of targeting errors. These must be sufficiently small for a Pareto improvement to be possible. How small depends on the amount of surplus flowing to those in the market $\bar{g}$, as well as the size of the budget relative to the total amount of burdens created, which is summarized in $\bar{\Delta}$. As efficiency gains are larger, the "budget" for targeting errors goes up. As the funding gap is larger, the margin of error shrinks.

The proof of condition 1 is in the appendix. The basic idea is that any targeting scheme will create some winners and losers. The condition asks whether there is enough efficiency gains ηG in the market to cover all of these losses, if by happy coincidence *gains are distributed perfectly so as to offset losses net of the transfers*. This is why the condition is necessary, but not sufficient, for a Pareto improvement to be possible.

Taking the condition to data: This condition is designed to be empirically testable with information that may be feasible to obtain for important policies. Other alternative

conditions exist that require different information. The terms on the right hand side of the inequality in condition 1 are market averages and do not require individualized data. In particular, condition 1 makes no assumption about the distribution of gains g_i, except that all are nonnegative. Testing the necessary condition does not require information about how the gains are distributed. What is required is a measure of the initial welfare burdens from a policy c_i. In some cases, for example in the case of marginal tax increases on commodities described next, it is straightforward to estimate this distribution from available data.

The condition is relevant for any arbitrary transfer scheme $T(\mathbf{X}_i)$. The empirical portion of this paper will try to predict initial burdens c_i with a set of covariates $\mathbf{X}_i$. Such a prediction exercise can then be mapped into a targeted transfer scheme. For any proposed transfer scheme, one can calculate the average absolute errors and check if the necessary condition is met. If a regression approach that minimizes the size of absolute errors (median regression) fails to generate an average absolute error small enough to satisfy the condition, then it is concluded that no feasible (i.e., based on the available covariates) transfer scheme can achieve a Pareto improvement.

In sum, condition 1 illuminates how the design of a Pareto improving transfer scheme is inherently a prediction problem. If the variation in policy burdens can be predicted accurately enough with the set of variables included in $\mathbf{X}$, then a Pareto improvement might be possible. What is "accurate enough" depends on the size of surplus gains—where surplus gains from a policy intervention are small, either because the externality is small or because quantities are not very sensitive to price, the accuracy window will be tight.

To illustrate further and to link the theory directly to the empirical analysis, I now discuss an interpretation of this model when the policy in question is an externality-correcting tax. Before proceeding, note that there are related necessary conditions that can be derived that might be useful for other types of policies. I.e., in some cases, one might know the distribution of g_i and wish to derive a condition that is agnostic about the distribution of c_i.

2.1 Pareto transfers for externality-correcting taxes

I consider now the specific case where the policy action in question is a marginal increase in an externality-correcting tax. In that case, the initial policy burdens c_i are the consumer surplus losses from increased prices. For marginal taxes starting from zero, revenue raised will equal burdens $C = R$, and burdens will be equal to observed baseline consumption, which means that only data on initial consumption is needed to measure the distribution of c_i. The average efficiency gains $\bar{g}$ require only an estimate of the demand derivative for the good and an estimate of the marginal external damages.

A few additional details and assumptions are useful to explicate.

The economy: Consider an economy with a good q that causes a negative externality and a quasi-linear numeraire. The agents in the market have exogenous and potentially heterogeneous incomes and heterogeneous utilities over q. Consumers are "small"—they assume their actions have no impact on aggregate outcomes. Preferences and income result in demand curves written $q_i(p+t)$, where p is the market price and t is any tax levied on the consumption of the good. The supply of q is assumed to be perfectly elastic, so the full burden of any tax is thus borne by buyers. Demand for q_i is assumed to be nonnegative. Denote the baseline consumption of the good (demand when taxes are zero) as $\tilde{q}_i \equiv q(p)$.

The externality: The externality is global, homogenous and linear, so that the total social damages depend on only the aggregate consumption of q, which is written as $Q \equiv \sum_i q_i$. Marginal damages per unit of Q are the sum of marginal damages to each individual, denoted ϕ_i. The aggregate marginal harm of the externality is $\Phi \equiv \sum_i \phi_i$. The total externality in the economy is thus ΦQ. This notation assumes that all of the externality gains are in the market, but this is not essential.

The tax: In this setting, the first-best outcome can be achieved through a standard Pigouvian tax on consumption equal to $t = \Phi$. Rather than that tax, I model here the introduction of an infinitesimal positive tax, starting from zero. As discussed further below, this is conservative towards finding a potential Pareto improvement because it implies that $\Delta = 0$ (by the envelope theorem, $R = C$). The same steps used below can be followed for a "small" tax using standard triangle approximations.

The effects of the tax: The higher price of q causes each consumer to lose private surplus. These losses are the policy burdens c_i. For an infinitesimal tax, this loss is equal to baseline consumption, $c_i = \tilde{q}_i$, by Roy's identity.

The tax also raises revenue from each consumer, denoted r_i. For an infinitesimal tax, the revenue raised is equal to the burden, so that total revenue $R = C$.

Finally, the tax creates a welfare gain equal to $\Phi \sum_i q_i' \equiv \Phi Q'$, where q_i' is the derivative of demand with respect to price for each consumer and Q' is the aggregate demand derivative. Each consumer experiences gains from the externality reduction, denoted $g_i = -\phi_i Q'$.

The transfer function: Revenue is recycled in a transfer scheme T that is assumed to be lump-sum and based on a vector of covariates $\mathbf{X}_i$. The idea is that a transfer function will depend on some characteristics (age, income, etc.). The transfer function cannot be tailored to each individual, but instead can only be targeted based on those variables. I focus on the case where all revenue is recycled, so that $\sum_i T(\mathbf{X}_i) = R$. Alternatives are discussed below. In order to match the empirical application, I assume that $T(\mathbf{X}_i)$ is linear in parameters.

It is possible to compensate losers perfectly in this framework by simply rebating tax

revenue according to the initial expenditure levels. That is, if $\mathbf{X}$ includes baseline consumption itself, then perfect prediction is possible. But, assuming agents understand the transfer scheme, this will eliminate (or at least dampen) the incentive to correct the externality and is thus self defeating. In the exposition, the covariates in $\mathbf{X}$ are assumed to be exogenous, but I discuss in the next section how the framework can be modified to account for transfer-induced behavioral responses among the covariates.

2.2 Discussion of the modeling assumptions for the case of a Pigouvian tax

The model makes a number of assumptions to deliver a tidy result. Most are easy to relax and are biased against the paper's main findings. I discuss them here before proceeding to the empirical analysis. Larger extensions, including some cases where a Pareto improvement might be easier to obtain, are discussed in section 5.

Starting from a pre-existing tax: The derivation assumes a zero marginal tax. This ensures that the initial consumer burdens are exactly offset (on average) by the revenue raised ($C = R$). Suppose instead that a tax exists, but that the tax is below marginal damages so that an increase is efficiency enhancing. The analysis can proceed by simply reinterpreting the externality gain as the difference between marginal damages and the existing tax rate (i.e., the uncorrected portion of the externality). The only difference is that, in this case, a marginal increase in the tax will create initial consumer burdens that are in excess of revenue raised ($C > R$). This just raises the funding gap Δ, which makes achieving a Pareto improvement even more difficult.

A non-marginal tax: The derivation assumes a marginal tax. For a non-marginal tax increase, burdens will exceed cost, which will again increase the funding gap Δ, without otherwise changing the problem, thus making a Pareto improvement more difficult.

Heterogeneity in behavioral response: For a non-marginal tax or when the initial tax is above zero, the behavioral response (the own price demand elasticity for the good) will figure into the burden calculation (i.e., the envelope condition does not apply). Where demand derivatives are homogenous, no new information is required to conduct an empirical test in these cases. But, if demand derivatives are heterogeneous, then information about the joint distribution of baseline consumption and demand derivatives is needed to calculate the distribution of burdens. This complicates empirical tests of the condition, but just as significantly, it greatly magnifies the real world task of compensating losers. It will frequently be easier to measure baseline consumption than to also estimate behavioral responses. Where the initial burden of a policy is harder to measure, the task of identifying and compensating

losers will be even more difficult.

Incidence (partial equilibrium): The model assumes complete pass through of prices to consumers and assumes that welfare can focus on only consumers. Where producers of the good also bear the burden, pass-through estimates are needed to divide up gains and the planner must consider targeted allocations on both sides of the market. The problem is otherwise the same. (Note that heterogeneity in pass through (for an example of which, see Stolper (2018)) would further raise information requirements.)

Incidence (general equilibrium): Corrective taxes can create burdens through a variety of general equilibrium effects and other channels (Fullerton 2011). For example, a carbon tax will affect factor prices. These general equilibrium effects may be substantial and heterogenous across groups (see, e.g., Goulder, Hafstead, Kim, and Long 2018). The partial equilibrium focus in this paper is motivated in part by practicality, but to the extent that the exercise is motivated by political economy concerns, this limited partial equilibrium view is likely the important one. If voters are unable to anticipate general equilibrium incidence effects, it is plausible that they heavily discount them in forming their judgments about how a policy will affect them. The immediate impact on prices and any promised transfer scheme are likely the dominant considerations for the empirical applications.

Measuring burdens and gains: This last point segues to a broader point that the full benefits (as well as the costs) of any externality correcting tax are difficult to measure. This is a challenge for the econometrician, but the challenge is just as great for the planner and the voter, which reinforces the core point of this paper. When the planner cannot measure the benefits or costs, it is no more possible to control the final distributional effects. And, when benefits and costs are difficult for the voter to perceive, it will be difficult to convince everyone that they are in fact benefitting. (Of course, if they can be fooled, then perhaps the political problem can be solved through deception rather than targeted transfers.)

Additional revenue used: The model assumes that the revenue used for transfers is equal to the revenue increase created by the policy. If additional revenue is available that is costless to acquire, then this can be introduced into the model simply by changing Δ to account for this supplementary revenue. In this case, the condition holds as written.

Revenue, however, is not costless to acquire. In terms of the question at hand, either this revenue is being taken from someone else outside the market who would thus require compensation, thereby defeating the broader goal of a Pareto improvement; or the burdens of raising additional revenue comes from market participants, perhaps through some other form of taxation. In this case, the use of the supplementary instrument aids Pareto compensation only if the revenue raised from the supplementary instrument exceeds its private welfare costs; i.e., if the marginal cost of public funds is below 1. In general we assume the opposite,

which suggests supplementary taxation is unlikely to help.[4]

Distortionary transfers: The model is described assuming that the covariate vector $\mathbf{X}$ is fully exogenous. Relaxing this assumption implies that the transfer scheme may create behavioral responses to $\mathbf{X}$, which raises several issues. As an example, suppose that $T(\mathbf{X})$ is rising in income (this is the case in the empirical examples below). Then the transfer scheme acts like an income subsidy. This may have efficiency benefits that can be incorporated into the analysis because it reduces the pre-existing tax on labor supply. Note, however, that to the extent that the transfer function exactly offsets the tax burden across the income distribution none of these incentives are an issue. In that case, the tax and transfer combined leave unchanged labor supply incentives so labor supply is unchanged, as argued in Kaplow (2004).

When the transfer function does not offset the tax burden exactly, it may create incentives (distortionary, or beneficial). Empirically, the slope of the empirical transfer functions and the plausible range of price elasticities on the variables used in the analysis—income, family structure, age and state of residence—imply that these feedback channels are economically insignificant.

This is not necessarily true, however, were covariates to include close proxies for the externality that might be highly elastic. For example, suppose one included vehicle fuel economy and commuting distance in a transfer scheme predicting burdens from the gasoline tax. Those variables would likely have large coefficients in the predictive transfer scheme $T(\mathbf{X})$, and they are probably as responsive as gasoline consumption itself. Basing transfers on these variables would obviously erode the efficiency potential of the gasoline tax (assuming agents understand the incentives). Empirically, I experiment with variables available in the CEX and find that they do not add much predictive power, suggesting that the issue is moot in the empirical cases below. Theoretically, I discuss the design of a second-best transfer scheme that trades-off greater prediction against these distortions in section 5.4.

Mechanism design: The core problem in this setup is incomplete information. If the planner knew the full set of root factors, including taste, that determined baseline consumption, then the planner could design precisely targeted transfers so as to achieve full compensation. This begs the question of whether a mechanism design approach would not

[4]To see this, assume some instrument raises supplementary revenue from individuals equal to s_i with $1/N \times \sum_i s_i = \bar{s}$ at welfare cost w_i with $1/N \times \sum_i w_i = \bar{w}$. In terms of the model, the revenue from this additional instrument will just change Δ (the funding gap), so that $C - R + S = \Delta$. We can remain agnostic about the distribution of welfare costs, and will simply subtract them from the efficiency gains in the derivation, so that the net gain from the policy suite is $g_i - w_i$. Then, the inequality in condition 1 becomes: $1/N \sum_i |c_i - T(\mathbf{X}_i)| > (\bar{g} - \bar{w}) - (\bar{\Delta} - \bar{s})$. One immediately sees that the introduction of s and w makes the right-hand side of the condition smaller (harder to satisfy) as long as $\bar{w} > \bar{s}$, that is, as long as private costs exceed revenue raised (the marginal cost of public funds exceeds 1).

yield a more favorable outcome. That is, might a mechanism be created that would cause each agent to honestly reveal the unpredictable portion of their heterogeneous demand for the good?

Consider first the case where the tax rate is imposed and initial burdens are created, and the mechanism in question pertains only to the allocation of the revenue through the transfer scheme. This problem is just dolling out transfers in a zero-sum fashion, so implementation theory will not be able to create incentives for honest revelation.

Instead, a mechanism design approach would have to involve more than just the transfer, such as a scheme where there are alternative tax rates on the good as well as a transfer scheme. Mechanism design solutions tend to leave some efficiency on the table in the form of rents to some types. Moreover, the imposition of budget balance consistent with the setup here is typically constraining in such settings.

For example, suppose there was simply an opt-out option where agents could avoid the tax but would not be eligible for a transfer. As long as the transfer in those schemes is based on the same covariates $\mathbf{X}$, this scheme will necessarily create winners and losers in the same way as any scheme analyzed in the framework above. Thus, if it creates losers, those losers would opt out. If anyone opts out, this will reduce revenue, shrinking the transfers of the remaining people, which leads to an unraveling. One could support a pooling equilibrium (everyone opts in) by imposing a penalty on those who opt out, but this is just creating losers by another name.

With all of this discussion about the setup in mind, we are now ready to move to the empirical analysis, which begins with a discussion of the data.

3 Consumption data on externality-generating goods

This paper uses data from the interview portion of the Consumer Expenditure Survey (CEX), which is a nationally representative sample of U.S. households, from 1996 to 2016. The CEX defines a unit of observation as a consumer unit, which is a set of individuals who reside together and are either related by blood or marriage, or who make financial decisions together.

Interviews consist of retrospective questions that ask about the consumer unit's total expenditures on various items over the prior three months. Units are interviewed four times, once each quarter, but not all units complete all four rounds of interviews. For the analysis below, expenditure categories are averaged over however many interviews are completed by a consumer unit, and then scaled to represent annual consumption amounts.

Table 1 shows summary statistics on expenditures. Key for this paper is that there is

Table 1: Household Expenditure Statistics by Category

	Mean	Median	St. Dev	CV	Pct 0
Motor fuels	$1,820	$1,398	$1,716	0.9	9%
Electricity	$1,143	$984	$913	0.8	9%
Natural gas	$413	$162	$611	1.5	42%
Alcohol	$230	$14	$485	2.1	48%
Tobacco	$318	$0	$788	2.5	71%
All energy	$3,377	$2,933	$2,423	0.7	3%
All sin goods	$3,925	$3,411	$2,757	0.7	2%

Table shows annualized expenditures by category for all households in sample (N=197,668). Dollar amounts are in $2015. Statistics are weighted by survey sample weights. All energy sums motor fuels, electricity and natural gas. All sin goods includes all five individual categories summed. CV is the coefficient of variation. Pct 0 is the percentage of consumer units reporting zero expenditures in the category.

wide variability in the consumption of all variables. For example, average consumer unit expenditures on motor fuels is $1,820, but the standard deviation is nearly as large as the mean, at $1,716.

There are two important caveats to be kept in mind regarding the use of the CEX in this study. First, the analysis is concerned with variance and predictability of consumption levels across households. The survey response may mismeasure true consumption either because of sampling variability or because of inaccuracies in self-reported responses. For a discussion of CEX data quality, see Meyer, Mok, and Sullivan (2015). Throughout the paper, I winsorize all expenditure variables at 1% in order to trim the influence of outliers. Data quality remains a concern, but it turns out that the expenditure data reported here implies almost exactly the same mean gallons per year estimates as the 2009 National Household Travel Survey (NHTS), which provides some reassurance. Further comparisons of CEX data to data on gasoline consumption from the the data and home energy expenditures from the Residential Energy Consumption Survey are included in appendix B.

Second, the CEX reports expenditures, not quantities. Externality-correcting policies are typically specific taxes (per unit) not ad valorem. To model an ad valorem tax on a product, only the total expenditure is required. Corrective taxes, however, will often take the form of a specific (per unit) tax. For example, a carbon tax will raise the price of gasoline by a constant amount per gallon. Thus, to model the impact of a carbon tax on gasoline consumption, we need to estimate the gallons of gasoline consumed by a household, based

Table 2: Summary Statistics of Demographic Variables

	Mean	St. Dev.	Min	Max
Before-tax income ($2015)	59,224	61,678	-419,200	971,100
Consumer unit (CU) size	2.4	1.5	1	29
Persons < 18 in CU	0.62	1.1	0	14
Persons >64 in CU	0.28	0.6	0	8
Urban indicator	0.91	0.29	0	1
Reference person married	0.50	0.50	0	1
Year	2006	6.1	1996	2016

on their reported expenditure and prices.

For gasoline and diesel fuels, I use data from the Energy Information Administration (EIA) on the sales-weighted, tax-inclusive, retail price of all grades of each fuel type at the closest available geographic match to the consumer unit. That is, where the CEX identifies a consumer unit's metropolitan statistical area and the EIA has city-specific prices, the consumer unit is assigned prices in the past quarter that are the average EIA price for that city. In other cases, matches must be made at the state or PADD level.

For other goods, determining the price paid by consumers is more challenging. Consider alcohol. Prices will vary widely if a consumer unit is purchasing low cost beer or high-end Scotch. As a result, for goods other than motor fuels, I focus on predicting expenditures directly (rather than predicted tax burdens), which translates directly to taxes under an ad valorem tax, recognizing that this is not how a true Pigouvian tax would be designed.

The core empirical task in the paper is to determine the degree to which demographic variables that might plausibly be used in a transfer function are able to predict variation in expenditures across consumer units. Table 2 summarizes the key variables used for this purposes, which are measures of income, household size and location.

4 A gasoline tax creates losers

The primary empirical application of this paper is a gasoline tax, which is modeled here as an efficient carbon-correcting policy. This is an important policy in its own right, and also has advantages in terms of modeling and measurement with the CEX. The conceptual goal of this analysis is to analyze an optimally designed Pigouvian tax. I thus focus on the gasoline tax as a well-targeted policy for correcting carbon externalities, but I discuss the implications of other driving-related externalities in the robustness section below.

In this section, I calculate the relative magnitude of welfare gains as compared to revenue raised from a motor fuel tax, and then demonstrate the degree to which demographic variables can predict motor fuel consumption. Specifically, I model a small tax increase of 10 cents on motor fuels (both gasoline and diesel) under the assumption that the carbon externality from motor fuel consumption is not corrected at all prior to the tax. That is, I am interpreting existing gasoline and diesel taxes as having been motivated by considerations about the optimal way to raise revenue, irrespective of a carbon externality. These assumptions are designed to be conservative against my findings, as they will maximize the implied welfare gains from carbon taxation.

4.1 What are the carbon externality gains from motor fuel taxation?

As described in the model, the welfare gain from a small tax on gasoline will be equal to the change in gasoline consumption induced by the tax times the externality per gallon. I assume that in the long run a gasoline tax will be borne completely by consumers so that prices will rise by 10 cents per gallon.[5]

The gasoline demand literature typically estimates elasticities, so I translate the 10 cent gasoline hike into a percentage price change using the average retail gasoline price facing the consumer unit at the time of the survey in its geographic location. I then use a gasoline price elasticity of -0.4, which is interpreted as a long-run price elasticity, to translate this price change into a change in gallons of fuel consumed.[6] By its very nature, it is challenging to estimate the long-run price elasticity of gasoline. I experiment with alternative values below.[7]

[5]Existing studies find evidence of high pass through rates for state gasoline taxes, with many studies consistent with full pass through Chouinard and Perloff (2004, 2007); Doyle and Samphantharak (2008); Marion and Muehlegger (2011). Fewer studies consider the federal gas tax, perhaps because it has changed much less often, which impedes econometric investigation. (Chouinard and Perloff 2004) conclude that only half of a federal tax increase is borne by consumers. If true, it would be important to consider the incidence on U.S. households through the producer side in interpreting the estimates. I return to this issue when discussing the empirical results.

[6]Small and Van Dender (2007) estimate long-run elasticities closer to half this magnitude. Hughes, Knittel, and Sperling (2008) conclude that the elasticity has been declining over time, finding preferred estimates well below -0.4. Espey (1998) finds a range of estimates that extend well beyond -0.4 in magnitude, but this is based on a variety of studies with varying credibility of empirical strategy. There is some suggestion that demand might respond more to gasoline taxes than price variation (Davis and Kilian 2011; Li, Linn, and Muehlegger 2014), though these estimates, taken from monthly changes in consumption, may be due inflated estimates due to consumers pre-buying in anticipation of price changes (Coglianese, Davis, Kilian, and Stock 2017). This difference seems unlikely to persist in the long run.

[7]I assume a homogeneous elasticity. Simple back of the envelope calculations make clear that allowing for heterogeneity will have unimportant impacts on the qualitative results because the tax is small.

I use the EPA's conversion factor to determine the tons of carbon emitted per gallon of gasoline consumed (17.6 pounds per gallon / 2205 pounds per metric ton for E10, or 22.5 pounds per gallon / 2205 pounds per metric ton for diesel) and then multiply by $40 for the social cost of carbon.

All of the assumptions here are designed to be generous in favor of creating larger externality benefits, and using the global social cost of carbon is foremost in that generosity. Climate benefits are largely realized in the future, and the majority of benefits will be realized outside of the U.S. Indeed, the current administration advocates use of a domestic social cost of carbon ranging from $1 to $6 in rule making. Thus, while there is vigorous debate about the right estimate of the social cost of carbon, it is exceedingly likely that $40 per ton exaggerates the benefits that accrue to current U.S. drivers.

4.2 Externality gains are much smaller than the initial burden and revenue raised

Because I am modeling a small gasoline tax, the initial burden (loss of consumer surplus from the higher price) will be approximately equal to the revenue, both of which are simply the price increase times the number of gallons of gasoline consumed by the consumer unit. But, to be more precise, I use the elasticity estimate to calculate the final quantity consumed, and use that to calculate revenue. The welfare loss is calculated using a linear approximation. Specifically, revenue raised from each household is equal to 10 cents times the new consumption level, which is equal to the current observed level of consumption (from data) minus the elasticity (-0.4) times the implied change in price (current price plus 10 cents divided by the current price, all minus 1). The initial private welfare loss is calculated as the new consumption level (as described above), plus the triangle, which is the change in consumption (as described above) times 1/2 times the tax (10 cents).

Table 3 shows these calculations for the estimation sample. The externality gains are $8.3 per consumer unit per year on average, while the revenue raised is $90 per consumer unit per year. Average costs imposed on consumers is slightly higher, at $91. The revenue raised is an order of magnitude larger than the externality gain. This has an important implication for the ability of the planner to create a Pareto improvement because, as shown by the theory, the externality gains represent the "error budget" available. A large amount of revenue needs to be reallocated via a transfer function, and the error budget is small relative to the revenue raised.

Table 3: Summary of the Impact of a 10-cent Gasoline Tax

	Mean	Standard deviation
Annual gallons consumed	926	75
Price change	6%	4%
Change in gallons	-26	32
Initial burden (c)	\$91	\$75
Net revenue (r)	\$90	\$74
Externality gain (g)	\$8.3	\$10

Table summarizes the impact on private welfare, the externality and revenue of a 10 cent gasoline tax, assuming an elasticity of -0.4.

4.3 Most variation in burdens is not predictable

The key suggestion of the theoretical model is that the degree to which the initial (pre-transfer) burden of the corrective tax can be predicted by variables that are used in the transfer function will determine whether a Pareto improvement is technologically feasible. Simple regression of the household level burden on variables that constitute the transfer function thus provides the required estimates. Below, I present results where the left-hand side variable is the estimated household level initial burden of a 10 cent gas tax.[8] All values are inflation adjusted to 2015.

The theory involves non-squared errors, so I present least absolute deviation (LAD) regressions that will minimize non-squared errors. But, I also present parallel specifications from OLS because the properties of OLS and the R^2 goodness of fit statistic is most familiar. Note that LAD will, by definition, yield lower absolute errors, but OLS, by definition, will maximize the R^2.

Table 4 presents the primary estimates from this exercise, with the top panel reporting OLS results. All regressions include year of sample fixed effects, which account for any time trends, though it turns out that excluding them has almost no impact on the results. Designing a transfer scheme that depends on any variables that are not strictly exogenous will create distortionary incentives. As a result, I focus attention first on the "most exogenous" variables that are likely components of a tax scheme, which are demographic indicators for household structure and geographic indicators for state and urban versus rural. Specifically, regressions include state dummies, an urban indicator, and dummy variables for the number of people in the household, as well as the number of minors, and the number over age 60.

[8] Because I am assuming a homogenous elasticity across households, this is equivalent to using initial baseline consumption (in gallons) as the left-hand side variable.

These variables predict just under 30% of the variation in gasoline tax burdens.

Column B adds a linear income control, followed by a non-parametric function of income (dummies in five-year bins) in column C. These provide a modest boost in the explanatory power, with the R^2 bumping up to .331 and .356, respectively. For reference, income by itself, without any demographic or geographic variables, explains only about 15% of variation (results not shown). Column C is my preferred specification. It is based on characteristics that are already part of the tax system, and could plausibly be used to design a tax reform or transfer scheme that accompanies an externality-correcting tax.

The unexplained variation in this specification is far too large to achieve a Pareto improvement. The average absolute error allows for direct comparison with the welfare gains from the externality. The residuals are around $45 per household. This compares to the $8.25 welfare gain. This is directly related to Condition 1: as long as the absolute average error exceeds twice the welfare gain, a Pareto improvement is not possible. Moreover, it is not just a matter of a few people being left as net losers. The best fitted scheme leaves more than one-third of households as net losers, even with the generous assumptions employed throughout.

Column D adds some clearly endogenous variables that would create significant distortions and are thus likely problematic variables for inclusion in a transfer scheme, including home energy consumption and dummies for the number of vehicles owned by the household, and dummies for the number leased. These variables do provide an additional boost to explanatory power, but even with vehicle ownership variables included, the variables explain less than half of the variation.

The bottom panel of table 4 shows LAD specifications. As expected, these lower the absolute error for identical specifications, but only by a very small amount.

Figure 1 shows the distribution of net losses, accounting for both the externality gain and the targeted transfers, based on column C in Table 4. A full 37% of households remain as net losers under this scheme.

For comparison, the figure also shows the distribution of net losses under a scheme where all households are rebated an equal share of the revenue. A similar fraction of households are net losers under both of these scenarios, but targeting radically reshapes the distribution.

The variables chosen here are the ones that are most likely to be used for a transfer scheme that operates through the tax code. The tax code is essentially a function of income and demographic structure of the household. As such, I interpret the results of Table 4 as demonstrating that gasoline expenditures are not predicted well enough to come remotely close to enabling a Pareto improvement. A Pareto improvement is not feasible.

It is worth restating the nature of the prediction dilemma at this point. Given informa-

Table 4: Predictability of Burden of a 10-cent Gasoline Tax

OLS	A	B	C	D
Avg. Abs. Error	\$46.6	\$45.0	\$44.2	\$39.9
R^2	.292	.331	.356	.456
LAD	E	F	G	H
Avg. Abs. Error	\$45.7	\$44.1	\$43.2	\$38.8
Pseudo-R^2	.181	.210	.226	.306
N	197,668	197,668	197,668	197,668
Year FE	Y	Y	Y	Y
Demo & geo controls	Y	Y	Y	Y
Linear income		Y	Y	Y
Binned income			Y	Y
Vehicles & energy				Y

Each letter represents a unique regression predicting the initial burden from a 10 cent gasoline tax. A and E include year fixed effects and dummy variables for number of household members, reference person married, number in household over 64, number under 18. B and F add a linear control for before tax household income. C and G add dummies for every \$5,000 of income. D and H add dummies for the number of vehicles owned or leased and level variables of expenditures on natural gas, electricity and heating oil.

Figure 1: Net Loss from 10-cent Gasoline Tax with Targeted Transfer

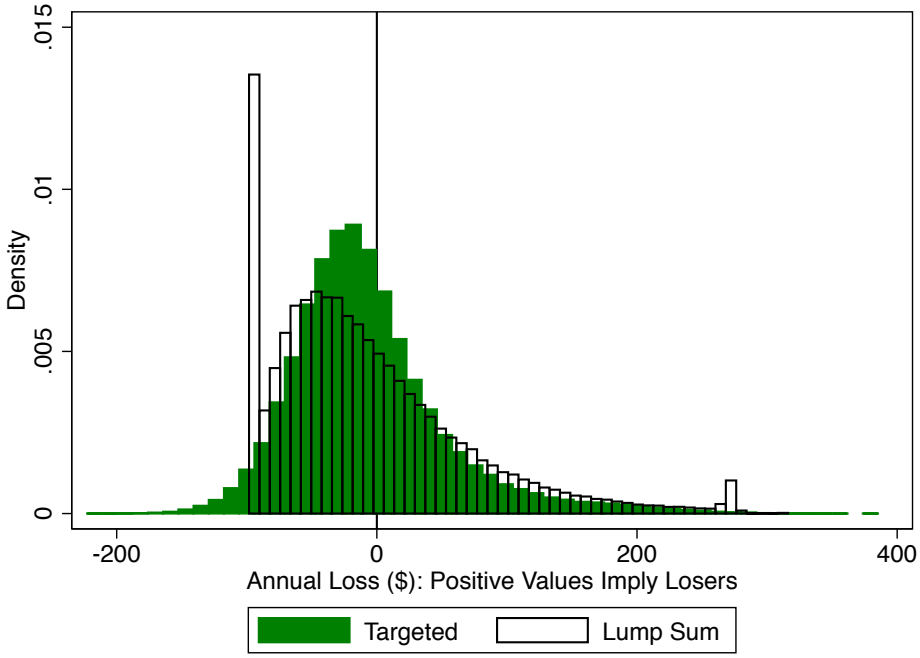

Figure shows the distribution of net impacts of a 10-cent gasoline tax, in dollars per year. A positive value implies a welfare loss. Results for equal per household transfer in transparent. Solid green indicates results for transfer scheme based on specification C in Table 4. The net impact is the private welfare loss, net of the targeted transfer scheme, net of the externality gain, which is assumed to be equal to each household.

tion directly on baseline fuel consumption, the planner could simply rebate every household exactly the burden imposed on it. But, if households understand this, then it completely (or at least significantly) undoes the price incentive—gasoline is not more expensive because the tax increase is rebated back, so there will be no externality gain. The thought experiment here is whether exogenous variables, like demographics and location of residence, are sufficient predictors. Of course even these variables are manipulable over time and not truly exogenous. I say more about how to think about that issue and how to incorporate intermediate variables (things that are likely responsive to a transfer scheme but are not gasoline expenditures itself) in section 5.4.

4.4 Machine learning marginally improves prediction

The problem posited here is fundamentally a prediction problem. It is thus a natural application for machine learning. A simple version is pursued here to see if initial steps can dramatically improve prediction.

Table 5 reports results of lasso regressions that predict the variation in tax burdens. The

Table 5: Lasso Regressions on Burden of 10-cent Gasoline Tax

	OLS (C)	Lasso	Lasso
Avg. Abs. Error	\$44.16	\$44.23	\$43.16
R^2	.356	.353	.379
Vars. Supplied		166	3,352
Vars. Selected		135	1,855
N	197,668	197,668	197,668
Year FE	Y	Y	Y
Demog. & geog. controls	Y	Y	Y
Linear income	Y	Y	Y
Binned income	Y	Y	Y
Additional interactions			Y

The first column repeats the OLS regression from Column C of Table 4. The second column runs a lasso regression on the same right hand side variable to perform a check for overfitting. The third column runs lasso with a large set of additional interactions. See text for details.

first column repeats column C from Table 4 for reference. The second column reports results from a lasso regression on the same variables to check for overfitting in the main specification. The specification uses a 10-fold cross validation and experiments with a range of lasso penalty parameters. Results suggest minimal overfitting. Lasso chooses a zero coefficient on 29 out of 166 variables, but this results in economically insignificant changes to prediction accuracy.

The third column introduces several thousand additional variables and uses the same 10-fold cross validation to select variables for inclusion, with the lasso penalty parameter chosen endogenously by the optimizer. Because the main specification includes predominantly binary dummy variables, the focus is on interactions, rather than higher order polynomials. The third column includes interactions of every income category and income linearly with year, state dummies, urban indicator, family size dummies, number in household under 18 dummies, number in household over 64 dummies, and a dummy for marital status. State by year fixed effects are also included. Despite selecting over 1,800 variables for inclusion, the improvement in prediction is minimal, and, from the point of view of achieving a Pareto improvement, barely perceptible. Additional experimentation with other interactions of these core variables produced similar results.

This is only a basic attempt to introduce prediction methods, but the lack of significant improvement from broader specification searches suggests that the variation in the burdens in the CEX is not predictable with the set of cross-sectional measures (household demographics, location of residence, and income) that is most plausibly usable as part of the tax code.

Table 6: Fraction of Losers Under Alternative Assumptions

Elasticity	Externality per gallon	Percent Winsorized	Percent Losers
-0.4	\$0.31	1%	37.0%
-0.4	\$2	1%	15.5%
-0.6	\$2	1%	9.7%
-0.8	\$2	1%	6.2%
-0.8	\$2	10%	3.0%

Each row comes from a separate regression of the burden of a 10-cent gasoline tax on the same set of covariates as specification C in Table 4. Each row varies a parameter as listed in the first three columns.

4.5 Robustness to parameter choices

In this section, I present results that alter three assumptions about the data. First and most simply, I increase the number of data points that are winsorized. Second, I modify the elasticity of gasoline consumption from -0.4 to -0.6 and then -0.8, to reflect higher estimates from the literature. Greater elasticities are important because they will lead to greater welfare gains, which aids the elimination of losers. Third, I greatly increase the externality per gallon of gasoline consumed, from around \$0.31 to \$2.

The higher latter number is based on accounting for non-greenhouse gas emissions from motor fuel consumption. Harrington, Parry, and Walls (2007) survey the literature and conclude that greenhouse gas emissions externalities are quite modest compared to accident and congestion externalities. A gas tax is a very poor instrument for targeting congestion, and a mediocre at best instrument for targeting accidents or local air pollution. Nevertheless, I now show cases where the gas tax could have much larger benefits in order to compare results.

In arriving at a \$2 per gallon externality, I modify the values from Harrington, Parry, and Walls (2007) to account for a higher accident externality, at \$0.91 per gallon based on Anderson and Auffhammer (2014), but interpret the carbon benefits as negligible. I then subtract off the sales-weighted average gas tax in the US of \$0.48. In terms of the literature on second-best gasoline taxes, however, note that this is still a generous interpretation in that it ignores fiscal interactions that exacerbate labor market distortions. Parry and Small (2005), for instance, argue that the second-best tax is only around 60% of marginal damages due to fiscal interactions.

Table 6 uses targeted transfers from specification C from Table 4, under alternative assumptions, to calculate the number of households that are net losers. Dramatically increasing

the interpreted externality gain per mile roughly halves the number of households who are net losers from a gasoline tax. Increase the elasticity of gasoline consumption to much higher rates further drives down the fraction of losers. In this scenario, the number of net losers is driven down to 6%. This is a modest number, but it should be kept in mind that there are many generous assumptions deployed in this case, so it should be interpreted as a frontier possibility rather than a realistic point estimate. Even in this case, some households are net losers. Finally, taking all of the prior assumptions and also winsorizing a full 10% of the data drives down the number of losers to 3%.

4.6 Other externality-correcting taxes are similar

The focus of this paper empirically is on a gasoline tax, but the CEX enables me to make quick assessment of the degree of predictability of other consumption categories that might be the focus of sin taxes. A gas tax has the advantage that it is relatively easy to translate expenditure data into quantities using gasoline price information, and hence to estimate the impact of a specific (per gallon) gasoline tax. The impact of other sin taxes is more difficult to determine because the goods are more heterogenous (e.g., there are many types of alcohol) and are subject to non-linear prices (e.g., two-part tariffs for electricity and natural gas).

Nevertheless, a broad picture of heterogeneity and predictability can be gained by simply regressing total expenditures in these categories on the demographic variables to see how much of the baseline expenditure variation is predictable. This exercise would exactly mimic the burden of an ad valorem sin tax, and they likely come close to mimicking the scale effect of sin tax levied per unit of the sin good in question.

Note that Table 1 shows that electricity has a similar coefficient of variation with motor fuels, but that other categories have even larger variability. OLS regressions in Table 7 shows the same pattern in terms of predictability. Electricity consumption is very similar in its predictability to motor fuels, but other sin goods are substantially harder to predict.

This analysis is incomplete, as it does not account for the welfare gains and is based on an ad hoc assumption about how a corrective tax would impact prices. But, the results suggest that a gasoline tax is likely the easiest place to achieve broad gains, and that the other externality-creating goods are likely to create even larger numbers of losers because of the greater inability to predict variation in baseline expenditures.

Table 7: Predictability of Other Sin Expenditures (OLS)

All statistics are R^2	A	B	C
Motor Fuels	.336	.382	.403
Electricity	.281	.324	.327
Natural gas	.179	.211	.214
Alcohol	.051	.126	.129
Tobacco	.043	.046	.050
All energy	.399	.471	.490
All sin goods	.367	.441	.459
N	197,668	197,668	197,668
Year FE	Y	Y	Y
Demog. & geog. controls	Y	Y	Y
Linear income		Y	Y
Binned income			Y

Each entry in the table is the R^2 from a separate regression that predicts expenditures (not burdens) on the category listed in the row, with control variables that vary by column. Column A includes year fixed effects and dummy variables for number of household members, reference person married, number in household over 64, number under 18. Column B adds a linear control for before tax household income. Column C adds dummies for every $5,000 of income.

5 Constructive next steps

The thesis of this paper is fundamentally negative. Not all losers can be compensated. This is an important observation, but it is also an unsatisfying place to stop. Several questions suggest themselves as next steps. I explore a few in this section, beginning with a discussion of several situations in which a Pareto improvement might be possible.

5.1 When might a Pareto improvement be achievable?

Benefits taxes: Intuitively, it should be easier to create a Pareto improvement where the welfare gains g_i are tightly correlated with the burdens c_i. Benefits taxes, for example, are intended to have this feature. Relatedly, Hall (2018) argues that congestion pricing can create a Pareto improvement. This possibility is due in part to the fact that the efficiency gains are closely tied to burdens; those paying a toll are paying directly for the benefit of reduced congestion.[9] Similarly, alcohol, cigarettes and sugary beverages are goods that may lead to externalities, but are often assumed to also be the source of internalities, such that the welfare improvements are concentrated among the heaviest users of the products.[10]

A close correlation between burdens and efficiency gains will imply that the net distribution of costs is less dispersed, which will compress the distribution of losses. Even so, Condition 1 is derived for a generic distribution of gains g_i, so the result holds regardless of how tight is the correlation between c_i and g_i.

Historical baselines: Another approach is to try to use historical baseline consumption to form the transfer scheme, which is the normal method in pollution permit allocations for firms (Schmalensee and Stavins 2017). This is harder to envision for households, though not impossible. It must accommodate entry and exit of households into the economy. It must also maintain a credible initial baseline that is not updated.

Broader policy packages: The focus of this paper is on considering a single corrective tax and the transfers that can be created with the revenue raised. If such a policy struggles to create a Pareto improvement, it begs the question of whether a broader set of policies considered together can yield a different result (i.e., fewer losers). This is related to the issue of logrolling. Here the question is whether multiple taxes taken together may create a less diffuse (or more predictable) distribution of burdens, which will occur for corrective taxes if

[9] The Pareto result in that paper hinges, however, on a claim that congestion actually reduces total throughput, which engenders a partial pooling equilibrium where some can opt out without leading to unraveling.

[10] Allcott, Lockwood, and Taubinsky (2018) analyze optimal corrective tax policy in the presence of heterogeneous internalities. The internality argument is sometimes said to hold for energy-consuming goods if consumers undervalue efficiency, though the evidence of significant behavioral biases is not clear (Allcott and Greenstone 2012).

the consumption of various goods are negatively correlated across agents. The fact that the combined raw energy expenditures are more predictable than each of the categories taken separately hints that this possibility may have empirical relevance.

Distributional implications of tax reductions or public expenditures: Another issue is that the revenue could be spent on public goods and services, rather than returned lump sum. Or, it could be used to lower particular taxes. These alternatives would further alter the distribution of burdens. It could be that this creates greater concentration of burdens, but it seems more likely that it creates further dispersion, as found in Cronin, Fullerton, and Sexton (2019) for the case of carbon taxes, thereby only making the problem harder.

5.2 Characterizing the trade-off between losses and revenue

This paper focuses on schemes where the budget available for transfers to compensate losers is equal to the revenue raised from the tax. The motivation is that this clearly isolates the possibility of a Pareto improvement, as compared to an identical economy with no corrective tax. In reality, there is no requirement that transfers use all of the revenue, nor is there a prohibition against using extra revenue taken from general funds.

In fact, efficiency dictates that all revenue from a corrective tax should be used to lower preexisting distortionary taxes (e.g., see Goulder 1995). Using revenue to compensate losers thus comes at a cost. As such, a policymaker might want to characterize the trade-off between loser compensation and revenue outlays. I describe one way of doing that here.

Specifically, I assume that the transfer function would be targeted so as to minimize typical losses for the case where outlays equal tax revenue, and that to scale outlays up or down, the transfer function for all households is scaled proportionally. I.e., consider an estimate of the targeting function, $T(\mathbf{X}_i)$ that would be used if all revenue were reallocated to consumers. Then write the total outlays as a fraction of revenue as $\theta = \sum_i T(\mathbf{X}_i)/R$. When $\theta = 1$, all revenue is spent on transfers. When $\theta = 2$, the outlay is double the revenue brought in by the tax.

I assume that individual transfers are all scaled proportionately, so that the transfer to consumer i is $\theta T(\mathbf{X}_i)$. Under this assumption, it is straightforward to characterize the number of losers, the average loss among losers, the variance in loss among losers, or other statistical moments that a decision maker might find useful in deciding how much revenue should be spent on compensation.[11]

[11] Proportional scaling may not be the optimal scheme, depending on the rationale for being concerned with losers. It is, however, an intuitive assumption and it is employed here to provide a tractable summary of information that a policymaker might use to make decisions. I discuss optimal transfers further below.

Figure 2: Fraction of Households Who are Net Losers As a Function of Outlay Ratio (θ)

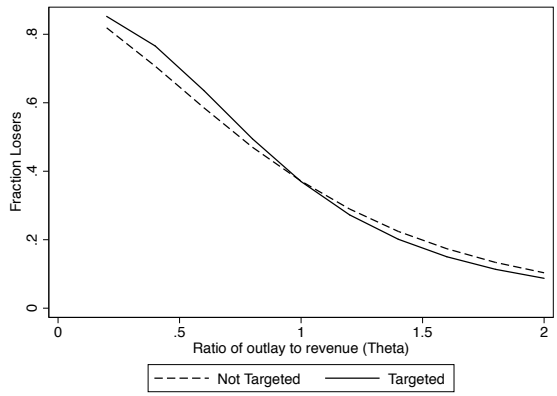

Figure shows fraction of households who are net losers as a function of the ratio of total transfers to revenue. Targeting is based on regression from column C, table 4.

Figure 3: Mean and Standard Deviation of Loss (Conditional on Losing) As a Function of Outlay Ratio (θ)

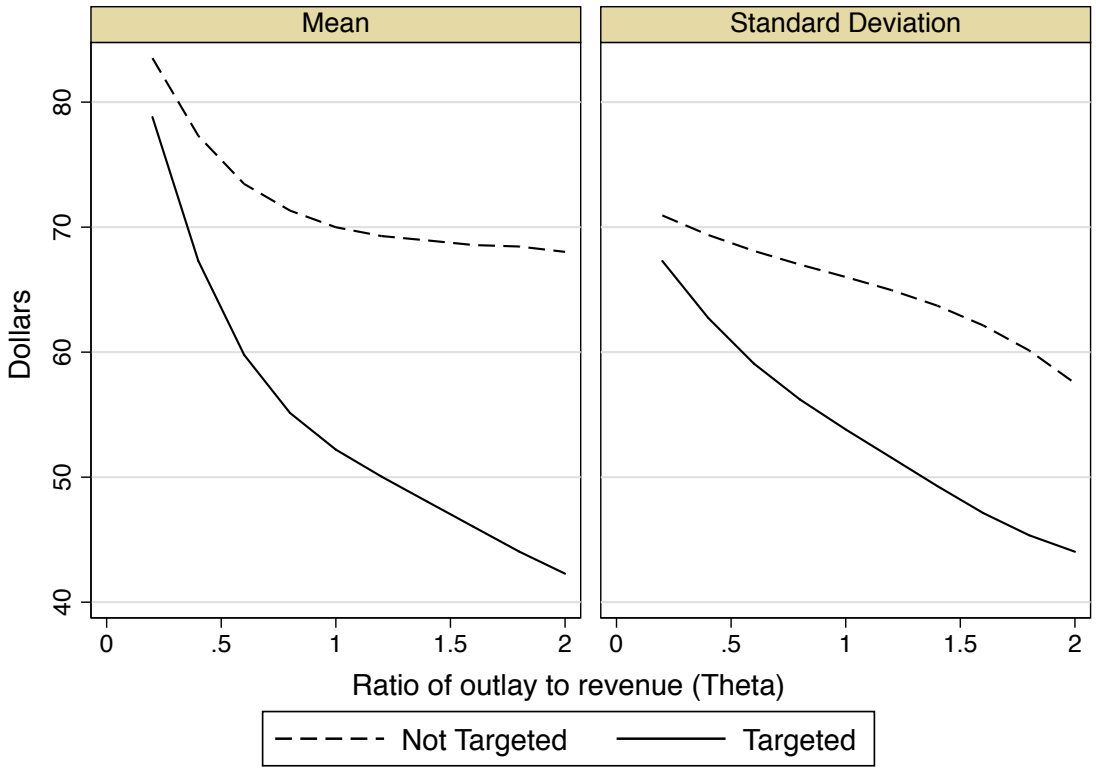

Figure shows mean and standard deviation of household losses as a function of the ratio of total transfers to revenue. Statistics are conditional on a household being a net loser. Targeting is based on regression from column C, table 4.

Figure 2 plots the fraction of households who are net losers from the ten-cent gas tax modeled above as a function of the targeted transfer scheme and the outlay ratio θ. The solid line shows results assuming that the targeting function is based on the predicted values from the specification in column C of table 4. For comparison, the dashed line shows the same fraction of losers under the assumption that all revenue is returned equally to each household. For either scheme, as expected, the fraction of losers sharply declines as outlays increase. Interestingly, there is little difference between the fraction who are losers between the targeted and untargeted schemes.

The value of targeting is more readily apparent when looking at the distribution of losses among losers as a function of revenue, which is plotted in figure 3. The left panel shows the average loss (conditional on a household being a net loser). Average losses among losers decline significantly as outlays increase, and they are much lower under targeting. The same is true for the standard deviation in losses (conditional on a household being a net loser), which is shown in the right panel.

Comparing the dashed and solid lines illustrates to the policymaker the value of targeting, and the slopes of the lines capture the trade-off between valuable revenue and compensating losers. For a decision maker concerned with achieving some degree of compensation, these types of statistics can convey valuable information. To fully evaluate alternative transfer schemes and decide how much revenue is worth dedicating to compensation, one requires a model of optimal loser compensation, which I turn to next.

5.3 Towards a politically optimized transfer scheme

This paper is fundamentally an exploration of how targeted transfers can alter the political prospects of efficiency-enhancing policies. The point of this paper is that one easy and appealing political solution—to say that everyone gains—will often be infeasible. Instead, transfer schemes will create winners and losers.

A next step would be to describe the "politically optimal" transfer allocation to accompany a Pigouvian tax (i.e., what is the transfer scheme that maximizes political support for a given tax?). A full investigation of this question is beyond the scope of his paper, but I make a few initial observations here.

Political targeting: First, note that a targeting scheme is well suited to the task of neutralizing political blocs. If a targeting scheme is based on predicted damages, the inclusion of any variable in the targeting equation would ensure that consumers with that characteristic are not losers on average. Thus, to the extent that a group of voters or stakeholders are deemed critical to the political survival of a policy, putting that variable into the transfer

prediction equation immediately creates a balance between winners and losers.

Perhaps the most obvious example is geography. If a transfer function, for example, includes state dummy variables, then winners and losers will not be concentrated among any state. More precisely, the average residual within each state will be zero, so that, for example, no senators would have constituents who lose on average.[12]

Alternative loss functions: If the goal was to minimize the number of losers, a planner would begin with a different loss function (neither OLS nor LAD). The LAD loss function is the correct one for checking a necessary condition for a Pareto improvement, but if the true motivation for the exercise is a political economy one, then it might be the case that the planner wishes to limit the number of losers to some politically acceptable number or more generally to control the distribution of losses.

A loss function that minimizes the number of losers is easy to program mathematically, but it will have impractical properties. A loser minimizing program would likely take the richest person in the sample and take all of their money in the form of a negative lump sum transfer so as to enhance the budget available for others. Some other restrictions are needed to yield useful results.

As a suggestive next step, I explore two alternative loss functions and show how optimizing against them changes the final distribution of burdens, as compared to an OLS benchmark. One way of capturing the notion of a desire to minimize losses, as opposed to simply accurately predict damages, is to introduce an asymmetry in the loss function. For example, a planner might not care at all about winners, but instead cares only about minimizing losers, but with a quadratic loss function for losses. Mathematically, this example is expressed with the following objective function, where the revenue constraint is included:

$$\min_{T} \sum_{i} \max(0, c_i - T(\mathbf{X}_i))^2 \quad \text{s.t.} \quad \sum_{i} T(\mathbf{X}_i) \leq R. \tag{1}$$

The optimal linear in parameters transfer function for expression 1 can be solved numerically. Note that even though the objective function does not value minimizing gains, the budget constraint implies a penalty for gains, so the results may not differ dramatically from OLS. Results are displayed in Figure 4 which plots the distribution of net gains for two policies that satisfy the same revenue constraint and target based on the same set of covariates (those used in column C of Table 4). The green histogram represents the distribution of net losses produced by OLS, where the white histogram represents the distribution of net losses

[12]This statement relates to the case where revenue outlay is equal to initial private burden, ignoring the distribution of gains. If gains are concentrated among groups, the prediction equation can be done on estimated net burdens to restore the result.

Figure 4: Distribution of Net Losses for Symmetric and Asymmetric Loss Functions

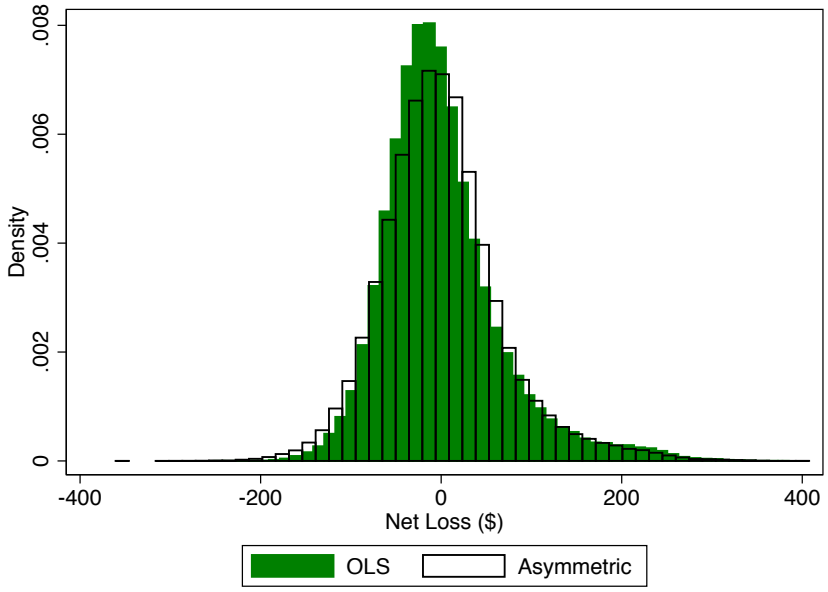

Figure shows distribution of net losses (positive vales) and gains (negative values) for baseline symmetric loss function (OLS), and for the asymmetric loss function in expression 1. Targeting is based on same covariates that are included in column C, table 4.

from the asymmetric loss function.

The differences are subtle. The asymmetric loss function somewhat reduces the right tail (the most extreme losers), and has a less peaked distribution. Even so, a regression of one set of residuals on the other produces an R^2 of 0.95 with a slope very close to one, suggesting that differences in the final outcome are small.

A second way of modifying the loss function to care more about losers is to change the exponent on the loss function.Median regression is less sensitive to outliers than is OLS. Here, we might be interested in being *more* sensitive to outliers, so that the transfer scheme is skewed more towards attempting to "reach" the biggest losers.

A parsimonious way to capture this idea is to specify a class of objective functions that minimize the absolute value of residuals raised to a power, denoted ρ:

$$\min_T \sum_i |c_i - T(\mathbf{X}_i)|^\rho \quad \text{s.t.} \quad \sum_i T(\mathbf{X}_i) \leq R. \tag{2}$$

This loss function nests OLS ($\rho = 2$) and median regression ($\rho = 1$). When $\rho > 2$, the loss function will put more weight on reducing the extreme outcomes, as compared to OLS.

To explore the sensitivity of final outcomes to this alternative objective function, I es-

Figure 5: Distributional Statistics Among Losers with Different Loss Function Exponents (ρ)

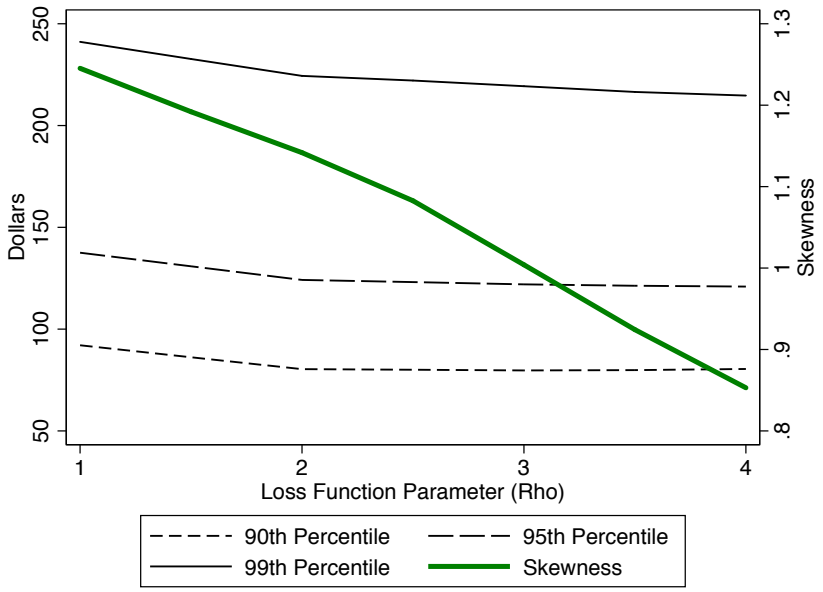

Figure shows percentiles characterizing the distribution of losers and skewness, as a function of the exponent on the loss function from expression 2. Targeting is based on same covariates that are included in column C, table 4.

timate a series of regressions for values of ρ ranging from 1 to 4, using a common set of covariates (those used in column C of table 4). As in the case of the asymmetric loss function, the distribution of net losses that emerges from numerical optimization differs only modestly across specifications. The distribution of losses is right skewed, and the most notable change across specifications is in the right tail (the biggest losers).

Figure 5 summarizes the impact on the right tail of the distribution by plotting the 90th, 95th and 99th percentiles of the net loss distribution, along with the skewness of the net loss distribution, as a function of ρ. These extreme data points decline as ρ increases, as expected, though the differences do not seem dramatic from an economic standpoint.

In brief, preliminary exploration of two alternative loss functions show the potential to think about optimizing the transfer scheme according to alternative criteria. They also, however, suggest that differences in the final distribution may be small. The magnitude of differences are due in part to the limited ability of the available covariates to predict consumption losses. Intuitively, with better predictability, the distribution of optimal net losses should differ more across specifications.

A more general model: Here I briefly describe a more general problem that posits some mapping between net outcomes and an individual's political support from a policy and

then asks how the transfer program would be designed to maximize the popularity of the policy.

Let the function PPI map net burdens $(c_i - T(\mathbf{X}_i))$ into support for the policy. Then, a planner who wishes to maximize aggregate (summed) support from the policy will solve:

$$\max_T \sum_i \alpha_i PPI(c_i - T(\mathbf{X}_i)) \quad \text{s.t.} \quad \sum_i T(\mathbf{X}_i) \leq R,$$

where α_i are the possibly heterogenous influence weights of each individual. With an estimated or calibrated PPI function, one can solve this problem in order to characterize the transfer scheme that maximizes popularity for the policy.[13]

With an estimated or calibrated PPI function, one can solve this problem numerically. The mapping PPI could be asymmetric in losses and gains, it might level out, and it might have particularly steep regions. Regardless, the setup suggests how this paper's key insight about prediction would carry forward—regardless of the exact shape of the PPI function, the planner's ability to accurately predict burdens with variables in the transfer function will determine the level of control the planner enjoys over the final distribution of net burdens. Better prediction will thus lead to higher values of the sum of the PPI function.

5.4 Second-best transfer schemes with endogenous covariates

Until now, this paper has not discussed efficiency trade-offs because the efficiency-enhancing policy (the Pigouvian tax) is taken as given and the transfer assumption is assumed to be based on exogenous covariates. Allowing for endogeneity in the covariates introduces trade-offs between loser compensation and efficiency, and it offers a distinct perspective on second-best transfer design. I discuss two versions of this trade-off, explained via the example of the gasoline tax.

Proxy targeting: The exercise in this paper is premised on the idea that a planner cannot simply return tax revenue back to households based directly on their consumption of gasoline because this would effectively undo the incentive to reduce consumption. That is, gasoline consumption itself cannot be a covariate in $\mathbf{X}$.

Even so, it might be possible to base the transfer scheme on inputs or close proxies that sacrifice some efficiency in externality mitigation (G shrinks) in exchange for improved targeting. I call this *proxy targeting*. For example, several factors determine a household's gasoline consumption, including the fuel economy of their vehicles, commuting distance,

[13] This is distinct from other ways of making the policy more politically successful, including altering the externality-correcting tax itself. It is also distinct from principal-agent problems in the policymaking process.

driving style and transit options. The efficiency appeal of a Pigouvian tax on gasoline is that it creates incentives for substitution along all of these margins.

The idea in proxy targeting is to condition transfers on some factors that determine the externality. This will dampen (or eliminate) the mitigation that comes through that margin and induce distortions along that margin, which sacrifices some efficiency gains, but perhaps in exchange for improved targeting. If some factor is highly correlated with consumption but not very responsive to a Pigouvian tax, it is a candidate factor. Commuting distance, for example, might explain a lot of the residual variation in consumption but represent a small part of the welfare gains, as it is probably quite inelastic.

In terms of achieving a Pareto improvement, one can see directly from Condition 1 that adding any variable that lowers G and lowers the average absolute error will make the condition easier to meet if and only if the reduction in the average welfare externality gain is less than half the reduction in the average absolute targeting error. Adding a proxy or input that has that characteristic would make achieving a Pareto improvement easier.

Orthogonal distortions: The other issue is the possible distortion of markets other than the externality-creating good itself. When the covariates in $\mathbf{X}$ are endogenous, the introduction of a transfer that is based on these covariates can induce distortions in the choice of $\mathbf{X}$.[14] In this discussion, I assume that any distortions in the $\mathbf{X}$ variables has no effect on G, and thus label this issue *orthogonal distortions*.

The framework can be generalized to take account of such distortions by positing that a planner solves a weighted combination of a targeting problem and the deadweight loss induced by behavioral distortions from the transfer scheme. Conceptually, a planner would solve an optimization problem of the form:

$$\min_{T} \sum_{i} |c_i - T(\mathbf{X}_i)| + \kappa DWL(T(\mathbf{X})) \quad \text{s.t.} \quad \sum_{i} T(\mathbf{X}_i) = R, \tag{3}$$

where DWL is the excess burden created by behavioral responses to tax rates, and κ is some scalar parameter that represents how much weight the planner puts on improved targeting versus excess burden. Intuitively, the transfer scheme will put larger taxes (or subsidies) on attributes in $\mathbf{X}$ as they are more valuable (conditional) predictors of c_i and as they induce less distortion.[15]

To gain more intuition, I make some highly simplifying assumptions that simplify the

[14] Empirically, the implicit taxes from the gasoline tax modeled above are small and the factors included are quite inelastic, suggesting that this is not likely a major factor in the empirical application as modeled.

[15] In practice, this will be similar to a standard welfare maximization scheme that introduces some form concern for horizontal equity (e.g., Auerbach and Hassett 2002) or that accounts for status quo allocations (e.g., Saez and Stantcheva 2016).

excess burden calculation. This transforms the problem in equation 3 into a combination of a well known tax problem and an empirical prediction problem that when combined map directly into the statistical problem of penalized regression.

Denote the covariates in $\mathbf{X}$ as $X_1, ..., X_j, ...X_J$. Above, these elements could include various nonlinear transformations of a variable of interest (e.g., income category fixed effects). Here, assume that each thing that is chosen by a household (e.g., income, number of children) is included only in a linear way in $T(\mathbf{X}_j)$. In this case, the slope of the transfer function with respect to variable j, denoted β_j, will be the implicit tax (or subsidy) to the good (or choice) j. Assume that $T(\mathbf{X})$ is linear in parameters and can be written as $T(\mathbf{X}_i) = \beta' X_i$, where β without a subscript is the vector of coefficients and β_j is one element.

Assume there are no pre-existing distortions (e.g., other taxes) on any X_j before the transfer scheme is introduced. Assume, for now, that there is no cross-price substitution between the variables in X_j. Then, a standard linear approximation of excess burden (DWL) indicates that the welfare cost of the implicit tax is $DWL_j \approx (\partial X_j / \partial \beta_j)\beta_j^2$, where the derivative here denotes an aggregate response across all i. Because there are no cross effects, total DWL will be the sum of DWL_j across j.

Lastly, I switch from a problem focused on minimizing absolute errors to one that minimizes squared errors, which might be consistent with the politically optimized scheme described in section 5.3. I make this switch here to allow a direct corollary with familiar properties of OLS. In this case, the planner's problem can be written as:

$$\min_{\beta} \sum_i (c_i - \beta' X_i)^2 + \kappa \sum_j \frac{\partial X_j}{\partial \beta_j} \beta_j^2 \quad \text{s.t.} \quad \sum_i T(\mathbf{X}_i) = R. \tag{4}$$

In this formulation, the planner chooses the transfer slopes (implicit taxes) β to minimize a weighted sum of the sum of squared residuals and deadweight loss. The first term in the objective function is the familiar OLS problem (minimizing the sum of squared residuals), and the second term is a the familiar problem of minimizing deadweight loss from optimal commodity tax theory, which has been studied since Ramsey (1927).

Empirically, the objective function in 4 is identical to the objective function used in penalized regression, like lasso and ridge. Here, the penalty parameters depend on the price-responsiveness of each variable, so this problem is just a form of ridge regression with covariate-differentiated penalties. (Standard statistical software, like glmnet in R, allow for such differentiation.) This suggests a method of numerically solving for second-best transfer schemes, and it offers a link between optimal tax models and regression problems that extends the observations in Jacobsen, Knittel, Sallee, and van Benthem (Forthcoming).

One can ask both which covariates j should be included in the transfer function, and what

should be their slopes. To answer the first question, suppose that all variables are entered with their OLS coefficients β^{OLS}. Adding a covariate j will reduce the sum of squared residuals by the partial-R^2 of j times the variance of c_i. It will increase deadweight loss by $\approx (\partial X_j/\partial_j^{OLS})(\beta_j^{OLS})^2$. Which is larger, accounting for the scale parameter κ, will determine whether variable j should be added.

The optimal slope coefficient, however, will not be simply the OLS coefficient. One can see intuitively from the problem (and by analogy to penalized regression) that as a variable is more elastic its coefficient will be attenuated away from the OLS benchmark. This is analogous to the familiar inverse elasticity rule in optimal commodity taxation. The core intuition is that variables that are more elastic will be penalized more, and their βs will be attenuated to reduce excess burden.

The discussion here makes a number of simplifications to produce some basic intuition. When there are pre-existing distortions, implicit taxes or subsidies can have different impacts on deadweight loss, including cases where a subsidy to a factor is welfare improving because it offsets pre-existing distortions. When the variables have cross-price effects, the impact on deadweight loss of a greater factor loading will have effects through related goods, making the problem analogous to the more general version of optimal Ramsey taxes on commodities. In addition, the discussion above abstracts from the fact that the targeting accuracy of a transfer function in fact changes as a result of changes in covariates. Thus, a proper solution to the empirical problem requires some fixed point in which the projected $\mathbf{X}$ vector is the equilibrium one. Numerical solutions are likely to be challenging, but the fact that this problem can be cast as a known form of penalized regression suggests a viable path.

6 Conclusion

This paper uses theory and data to argue that policies like Pigouvian taxes—which improve social efficiency but create heterogeneous costs and benefits—will inevitably create some losers because transfers targeting the losers will tend to be imprecise.

The theory demonstrates how one's ability to compensate losers depends on the predictability of heterogeneous policy burdens and the size of efficiency gains. The theory delivers a specific test that can be taken directly to data. Empirically, the case of a gasoline tax is considered, and the possibility of a Pareto improvement is soundly rejected. Preliminary evidence on other externality creating goods suggests the same conclusion. In short, Pigouvian taxes create losers.

This is an important conclusion as it suggests the need for nuance in a range of important policy debates. Economists sometimes argue that efficiency-enhancing policies, at least in

principle, can be paired with targeted transfers so as to rationalize completely abstracting from distributional implications and judging policies purely on efficiency grounds. This paper argues for more caution in this line of reasoning. The fact that a policy creates losers is not in and of itself a reason to reject the policy, but it does point out one reason why efficiency enhancing policies may not prevail in the policy-making process.

It is worth stating again that a concern with compensating losers is not born of the objective of maximizing social welfare. Standard social welfare maximization does not give any special consideration to the status quo allocation, and gains and losses per se are irrelevant. The informal motivation of this paper and its concern with compensating losers is about the political process. Pockets of losers who are particularly harmed by a policy may organize to obstruct it. If one takes the view that efficiency-enhancing policies are in fact desirable, then the fact that not all losers can be compensated should shift attention to the question of how many losers must be compensated, and by how much, in order to enable an efficient policy to prevail.

The final portion of the paper is intended as a first step in that direction. It demonstrates the value of targeting in compressing the distribution of winners and losers as a function of total revenue expended, and it experiments with alternative objective functions that aim to prevent especially large losses from occurring. A deeper exploration of these alternative targeting plans could further aid in the constructive design of policy packages that preserve economic efficiency while satisfying political constraints generated by distributions of burdens.

References

Akerlof, George A. 1978. "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review* 68 (1):8–19.

Allcott, Hunt and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives* 26 (1):3–28.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky. 2018. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." Manuscript: University of California.

Anderson, Michael and Maximilian Auffhammer. 2014. "Pounds that Kill: The External Costs of Vehicle Weight." *Review of Economic Studies* 82 (2):535–571.

Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1-2):55–75.

Auerbach, Alan J. and Kevin A. Hassett. 2002. "A New Measure of Horizontal Equity." *American Economic Review* 92 (4):1116–1125.

Bovenberg, A. Lans and Lawrence H. Goulder. 2001. "Neutralizing the Adverse Industry Impacts of CO2 Abatement Policies: What Does it Cost?" In *Behavioral and Distributional Effects of Environmental Policy*, edited by Carlo Carraro and Gilbert E. Metcalf. NBER, 45–90.

Bovenberg, A. Lans, Lawrence H. Goulder, and Derek J. Gurney. 2005. "Efficiency Costs of Meeting Industry-Distributional Constraints Under Environmental Permits and Taxes." *RAND Journal of Economics* 36 (4):951–971.

Bovenberg, A. Lans and F. van der Ploeg. 1994. "Environmental policy, public finance and the labour market in a second-best world." *Journal of Public Economics* 55:349–390.

Burtraw, Dallas and Karen Palmer. 2008. "Compensation Rules for Climate Policy in the Electricity Sector." *Journal of Policy Analysis and Management* 27 (4):819–847.

Burtraw, Dallas, Richard Sweeney, and Margaret Walls. 2008. "The Incidence of US Climate Policy: Where You Stand Depends on Where You Sit." Resources for the Future Discussion Paper.

Chouinard, Hayley and Jeffrery M. Perloff. 2004. "Incidence of Federal and State Gasoline Taxes." *Economics Letters* 83 (1):55–60.

Chouinard, Hayley H. and Jeffrery M. Perloff. 2007. "Gasoline Price Differences: Taxes, Pollution Regulations, Mergers, Market Power, and Market Conditions." *The B.E. Journal of Economic Analysis & Policy* 7 (1 (Contributions)):1–26.

Coady, David, Margaret Grosh, and John Hoddinott. 2004. *Targeting of Transfers in Developing Countries: Review of Lessons and Experience*. Washington, DC: World Bank.

Coase, R.H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3:1–44.

Coglianese, John, Lucas W. Davis, Lutz Kilian, and James Stock. 2017. "Anticipation, Tax Avoidance, and the Price Elasticity of Demand for Gasoline." *Journal of Applied Econometrics* 32 (1):1–15.

Cremer, Helmuth, Firouz Gahvari, and Norbert Ladoux. 1998. "Externalities and Optimal Taxation." *Journal of Public Economics* 70:343–364.

———. 2003. "Environmental Taxes with Heterogenous Consumers: An Application to Energy Consumption in France." *Journal of Public Economics* 87 (12):2791–2815.

Cronin, Julie Anne, Don Fullerton, and Steven E. Sexton. 2019. "Vertical and Horizontal Redistribution from a Carbon Tax and Rebate." *Journal of the Association of Environmental and Resource Economists* .

Davis, Lucas W. and Lutz Kilian. 2011. "Estimating the Effect of a Gasoline Tax on Carbon Emissions." *Journal of Applied Econometrics* 26:1187–1214.

Davis, Lucas W. and Christopher R. Knittel. 2019. "Are Fuel Economy Standards Regressive?" *Journal of the Association of Environmental and Resource Economists* .

Dinan, Terry. 2012. "Offsetting a Carbon Tax's Costs on Low-Income Households." Https://www.cbo.gov/publication/43713.

Doyle, Jr., Joseph J. and Krislert Samphantharak. 2008. "$2.00 Gas! Studying the Effects of a Gas Tax Moratorium." *Journal of Public Economics* 92 (3-4):869–884.

Espey, Molly. 1998. "Gasoline Demand Revisited: An International Meta-Analysis of Elasticities." *Energy Economics* 20 (3):273–295.

Fischer, Carolyn and Wiliam A. Pizer. 2019. "Equity versus Efficiency in Energy Regulation." *Journal of the Association of Environmental and Resource Economists* .

Fullerton, Don. 2011. "Six Distributional Effects of Environmental Policy." *Risk Analysis* 31 (6):923–929.

Gauthier, Stèphane and Guy Laroque. 2009. "Separability and Public Finance." *Journal of Public Economics* 93:1168–1174.

Goulder, Lawrence H. 1995. "Environmental Taxation and the Double Dividend: A Reader's Guid." *International Tax and Public Finance* 2 (2):157–183.

Goulder, Lawrence H., Marc A.C. Hafstead, and Michael Dworsky. 2010. "Impacts of Alternative Emissions Allowance Allocation Methods Under a Federal Cap-and-Trade Program." *Journal of Environmental Economics and Management* 60 (3):161–181.

Goulder, Lawrence H., Marc A.C. Hafstead, Gyu Rim Kim, and Xianling Long. 2018. "Impacts of a Carbon Tax Across US Household Income Groups: What are the Equity-Efficiency Trade-Offs?" NBER Working Paper 25181.

Grainger, Corbett A. and Charles D. Kolstad. 2010. "Who Pays a Price on Carbon?" *Environmental and Resource Economics* 46 (3):359–376.

Hall, Jonathan D. 2018. "Pareto Improvements from Lexus Lanes: The Effects of Pricing a Portion of the Lanes on Congested Highways." *Journal of Public Economics* 158:113–125.

Harrington, Winston, Ian Parry, and Margaret Walls. 2007. "Automobile Externalities and Policies." *Journal of Economic Literature* 45 (2):373–399.

Hassett, Kevin A., Aparna Mathur, and Gilbert E. Metcalf. 2009. "The Incidence of a U.S. Carbon Tax: A Lifetime and Regional Analysis." *Energy Journal* 30 (2):155–178.

Hughes, Jonathan E., Christopher R. Knittel, and Daniel Sperling. 2008. "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *The Energy Journal* 29 (1).

Ito, Koichiro and James M. Sallee. 2018. "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards." *Review of Economics and Statistics* 100:319–336.

Jacobs, Bas and Ruud A. de Mooij. 2015. "Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax." *Journal of Environmental Economics and Management* 71:90–108.

Jacobsen, Mark R., Christopher R. Knittel, James M. Sallee, and Arthur A. van Benthem. Forthcoming. "The Use of Regression Statistics to Analyze Imperfect Pricing Policies." *Journal of Political Economy* .

Kaplow, Louis. 1989. "Horizontal Equity: Measures in Search of a Principle." *National Tax Journal* 42 (2):139–154.

———. 2004. "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy." *Journal of Economic Perspectives* 18 (4):159–175.

———. 2012. "Optimal Control of Externalities in the Presence of Income Taxation." *International Economic Review* 53 (2):487–509.

Li, Shanjun, Joshua Linn, and Erich Muehlegger. 2014. "Gasoline Taxes and Consumer Behavior." *American Economic Journal: Economic Policy* 6 (4):302–342.

Marion, Justin and Erich Muehlegger. 2011. "Fuel Tax Incidence and Supply Conditions." *Journal of Public Economics* 95 (9-10):1202–1212.

Mathur, Aparna and Adele C. Morris. 2014. "Distributional Effects of a Carbon Tax in Broader US Fiscal Reform." *Energy Policy* 66:326–334.

Metcalf, Gilbert E. 2009. "A Distributional Analysis of Green Tax Reforms." *National Tax Journal* 52 (4):655–681.

———. 2018. *Paying for Pollution: Why a Carbon Tax is Good for America*. Oxford University Press.

Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29 (4):199–226.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2):175–208.

Musgrave, Richard. 1959. *Theory of Public Finance; A Study in Public Economy*. New York: McGraw-Hill.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

———. 1982. *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities*. Yale University Press.

Ostrom, Elinor. 1990. *Governing the Commons: THe Evolution of Institutions for Collective Action*. Cambridge University Press.

Parry, Ian W.H. and Kenneth A. Small. 2005. "Does Britain or the United States Have the Right Gasoline Tax?" *American Economic Review* 95 (4):1276–1289.

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

Pizer, Wiliam A. and Steven E. Sexton. 2019. "Distributional Impacts of Energy Taxes." *Review of Environmental Economics and Policy* 13 (1):104–123. Manuscript: Duke University.

Poterba, James M. 1991. "Is the Gasoline Tax Regressive?" *Tax Policy and the Economy* 5:145–164.

Ramsey, Frank. 1927. "A Contribution to the Theory of Taxation." *Economic Journal* 37 (145):47–61.

Rausch, Sebastian, Gilbert E. Metcalf, and John M. Reilly. 2011. "Distributional Impacts of Carbon Pricing: A General Equilibrium Approach with Micro-Data for Households." *Energy Economics* 33:S20–S33.

Saez, Emmanuel. 2002. "The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogenous Tastes." *Journal of Public Economics* 83 (2):217–230.

Saez, Emmanuel and Stefanie Stantcheva. 2016. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review* 106 (1):24–45.

Sandmo, Agnar. 1975. "Optimal Taxation in the Presence of Externalities." *Swedish Journal of Economics* 77 (1):86–98.

Schmalensee, Richard and Robert N. Stavins. 2017. "Lessons Learned from Three Decades of Experience with Cap and Trade." *Review of Environmental Economics and Policy* 11 (1):59–79.

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

Stolper, Samuel. 2018. "Local Pass-Through and the Regressivity of Taxes: Evidence from Automotive Fuel Markets." Manuscript: University of Michigan.

Varian, Hal R. 1994. "A Solution to the Problem of Externalities When Agents Are Well-Informed." *American Economic Review* 84 (5):1278–1293.

West, Sarah E. 2004. "Distributional Effects of Alternative Vehicle Pollution Control Policies." *Journal of Public Economics* 88 (3-4):735–757.

Williams, Roberton C., III, Hal Gordon, Dallas Burtraw, Jared C. Carbone, and Richard D. Morgenstern. 2015. "The Initial Incidence of a Carbon Tax Across Income Groups." *National Tax Journal* 68 (1).

A Appendix: Proofs

Condition 1. *Let c_i be the private burdens from a policy, N be the number of agents in the market, $T(\mathbf{X}_i)$ be a transfer scheme, $\bar{\Delta}$ the average funding gap, and $\bar{g}$ be the average efficiency gain accruing to those in the market. A Pareto improvement is not possible if the average absolute targeting errors exceed twice the average efficiency gain in the market minus the average funding gap; i.e., if*

$$\frac{1}{N}\sum_i |c_i - T(\mathbf{X}_i)| > 2\bar{g} - \bar{\Delta}$$

then a Pareto improvement is not possible.

The proof proceeds by stating the total size of all losses (initial burdens net of transfer, excluding the surplus gains) among losers as an algebraic expression of the average absolute error and the average funding gap. If the total losses among losers exceeds the total surplus gains enjoyed by participants in the market, then it is not possible that these surplus gains are distributed so as to compensate all losers.

For any transfer regime $T(\mathbf{X}_i)$ and distribution of costs c_i, partition the data into losers (anyone with $c_i - T(\mathbf{X}_i) > 0$) and winners (anyone with $c_i - T(\mathbf{X}_i) \leq 0$). Denote the set of losers as $i \in L$, with their number being N_L. Denote the set of winners as $i \in W$, with their number being N_W.

Having partitioned the data into winners and losers, one can define the total losses among losers (which will be positive by construction), denoted Z_L and the total losses (which will be negative by construction) among winners as Z_W:

$$Z_L \equiv \sum_{i \in L}(c_i - T(\mathbf{X}_i)) > 0 \qquad Z_W \equiv \sum_{i \in L}(c_i - T(\mathbf{X}_i)) < 0.$$

The sum Z_L is the total amount of loss among losers. If this loss exceeds the total efficiency gains $\sum_i g_i$, then it is impossible to achieve a Pareto improvement, because even if those gains are distributed in the most favorable way among all losers, there are not enough gains to compensate all losers. The goal now is to redefine Z_L in terms of the average absolute error among all i (because this relates to an empirically estimable object) and the average funding gap.

The funding gap is the amount by which initial costs exceed revenue available for transfers, $\Delta = R - C$, with the budget constraint implying that $\sum_i T(\mathbf{X}_i) = R$. The funding gap is equal to the sum of Z_L and Z_W: $\Delta = Z_W + Z_L$. (To see this: $\Delta = C - R = \sum_i (c_i - T(\mathbf{X}_i)) = Z_L + Z_W$, with the latter equality true because Z_L and Z_W just partition the full set.) We will want to relate absolute values and will want an expression to substitute out Z_W. Because Z_W is always negative and Z_L is always positive:

$$|Z_W| = |Z_L| - \Delta \tag{5}$$

For notational convenience, denote the average absolute error as $|c_i - T(\mathbf{X}_i)| = |\varepsilon|$. Denote

the average absolute error among all individuals by $|\bar{\epsilon}|$, with $|\bar{\epsilon_L}|$ and $|\bar{\epsilon_W}|$ representing the average absolute error of losers and winners respectively. The average absolute errors are equal to Z divided by N for each group:

$$|\bar{\epsilon_L}| = \sum_{i \in L} |c_i - T(\mathbf{X}_i)| = \frac{|Z_L|}{N_L} \quad \text{and} \quad |\bar{\epsilon_W}| = \sum_{i \in W} |c_i - T(\mathbf{X}_i)| = \frac{|Z_W|}{N_W}. \tag{6}$$

The average absolute value of all of the data is simply the sample-size weighted average of the absolute average among winners and losers, which is written in terms of Z_W and Z_L via substitution of (6):

$$|\bar{\epsilon}| = \frac{N_L|\bar{\epsilon_L}| + N_W|\bar{\epsilon_W}|}{N_L + N_W} = \frac{|Z_L| + |Z_W|}{N}.$$

Substituting for $|Z_W|$ using (5) yields:

$$|\bar{\epsilon}| = \frac{|Z_L| + |Z_W|}{N} = \frac{|Z_L| + |Z_L| - \Delta}{N}.$$

Rearrange to solve for an expression of $|Z_L|$, which is positive by construction:

$$Z_L = |Z_L| = \frac{N}{2}|\bar{\epsilon}| + \frac{\Delta}{2}.$$

This is an expression for the total loss among losers. If this loss exceeds the sum of efficiency gains, then a Pareto improvement is not possible. I.e., a Pareto improvement is not possible if $Z_L = \frac{N}{2}|\bar{\epsilon}| + \frac{\Delta}{2} > \sum_i g_i$. Rearranging yields the result:

$$\frac{N}{2}|\bar{\epsilon}| + \frac{\Delta}{2} > \sum_i g_i \Leftrightarrow |\bar{\epsilon}| > 2\bar{g} - \bar{\Delta}.$$

The statement in the proof uses the definition of $|\bar{\epsilon}|$. □

B Appendix: Data comparisons

The CEX was chosen as the primary data source for this analysis because it includes a rich set of demographic covariates and a measure of gasoline expenditures, and because it is the standard data source in the most closely related literature. Gasoline expenditures, however, are based on self-reports and may be subject to mismeasurement. If there is a lot of noise in the expenditure data, this will make prediction more difficult. This section attempts to establish some sense of the reliability of CEX data by comparison to other surveys.

Of course, at the very highest level, problems of measurement do not challenge the key thesis of this paper. Instead, these problems reinforce it. If the best available data on expenditures are noisy measures of true burdens, it only makes it more difficult to design an accurate targeting scheme and thereby to compensate losers.

The National Household Travel Survey

An alternative measure of motor fuel consumption can be taken from the National Household Travel Survey, which is a nationally representative survey performed most recently in 2001, 2009 and 2017. That survey gathers a measure of annual vehicle miles traveled and then divides by the EPA estimated fuel economy of a vehicle to arrive at an estimate of annual fuel consumed. This is multiplied by average gasoline prices from the Energy Information Administration to impute expenditures. In contrast, the CEX asks consumers directly about expenditures.

The 2009 version of the NHTS is the most recent survey in which the miles traveled variable was based on two odometer readings (the survey respondent is asked to look at their odometer), rather than a retrospective self report. I compare the motor fuels expenditure data from that survey to the CEX from 2009. Figure 7 shows that fuel expenditure distribution from the two surveys for different samples. The top panel shows all households. This shows that the NHTS has higher expenditures on average, with a substantially longer right tail.

Figure 6: Comparison of Distribution of Implied Fuel Expenditure in CEX and NHTS

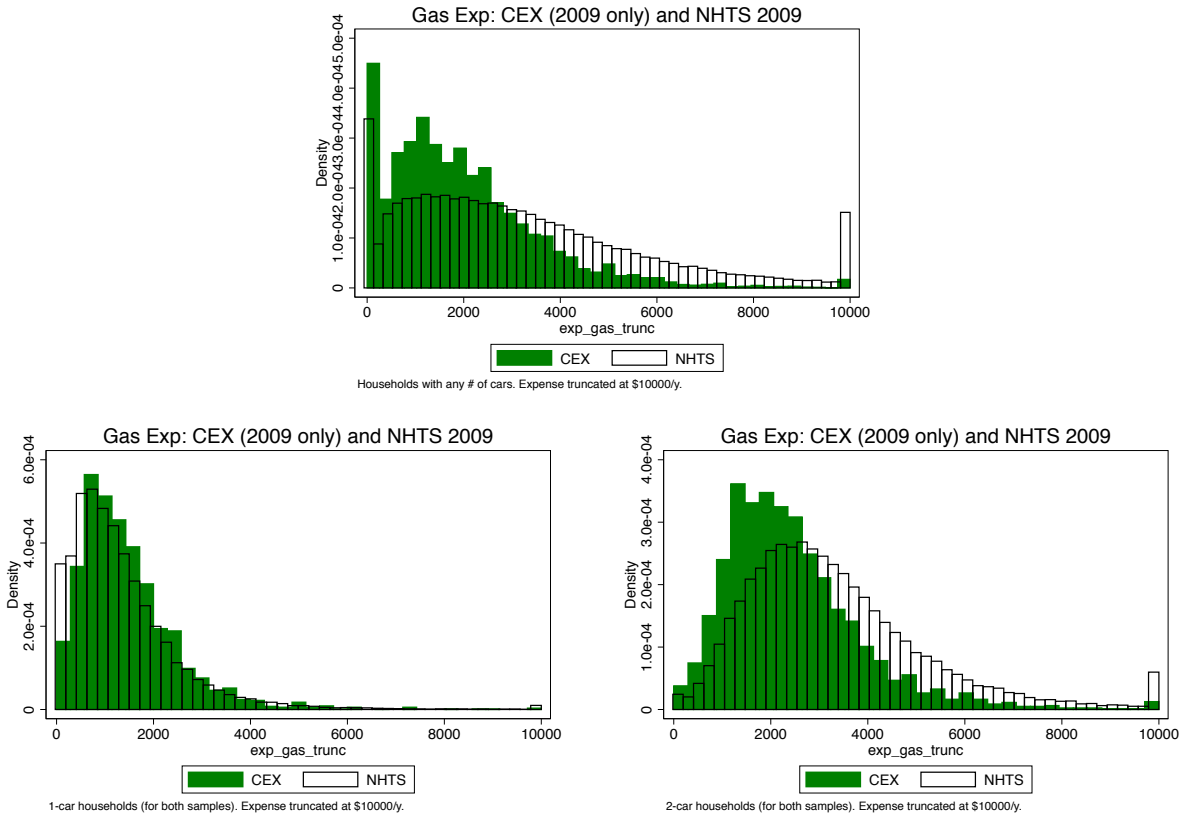

Figure shows histogram of estimated annual fuel expenditure by households using CEX and NHTS data. Left panel is for households with one vehicle. Right panel is for households with two vehicles. Both are from 2009 surveys. All distributions are truncated at $10,000 of annual expenditure.

In part this may be due to differences in unit definitions across the two surveys, as

Table 8: Predictability of Gasoline Expenditures in CEX versus NHTS

	CEX	NHTS	CEX	NHTS	CEX	NHTS	CEX	NHTS
R^2	.278	.232	.250	.267	.359	.380	.322	.431
Base controls	Y	Y	Y	Y	Y	Y	Y	Y
Number of vehicles	N	N	N	N	Y	Y	Y	Y
Weighted	N	N	Y	Y	N	N	Y	Y
N	9,116	137,938	9,116	137,938	9,116	137,938	9,116	137,938

Table compares 2009 CEX to 2009 NHTS. Dependent variable is annualized gasoline expenditures. Base controls include income, Census regions, urban dummy, family size and number of persons over 18. The additional variable is total number of vehicles in the household.

the CEX is broken into smaller consumer units than the household definition used in the NHTS. Differences persist, however, when comparing households with the same number of members. The bottom two panels of Figure 7 compare households with one vehicle (on the left) and with two vehicles (on the right). In particular for the one vehicle households, the distributions do fit better. Nevertheless, the two data sources do show non-trivial differences in this fundamental measure.

Though there are some advantages to the measure of fuel expenditure in the NHTS, it has the disadvantages of requiring imputation of fuel economy and gasoline prices. Gasoline prices vary significantly across locations and time. Fuel economy varies substantially with where a vehicle is driven. Thus, it is not obvious which survey measure is more reliable. Regardless, the fact that there are substantial differences suggests that mismeasurement could be important.

Ultimately for the purposes of this paper, what matters is predictability. To compare predictability across the surveys I identify a set of demographic variables that appear to be defined consistently in both surveys: income, Census region, an urban indicator, family size and number of persons over 18. Table 8 reports the R^2 for parallel regressions of gasoline expenditures on these controls, varying the set of controls and whether the regressions using sample weights.

For the base set of controls, the CEX and NHTS show very levels of predictability as summarized by the R^2. This is true regardless of weighting. In additional specifications (not shown), the results change very little when using dummies for the household size variables or adding state dummies instead of Census regions. The one difference that did emerge in a specification search was that the total number of vehicles owned by the household has a stronger explanatory power in the NHTS, and in particular when weighting, this variable notably increases prediction accuracy. The NHTS collects mileage information (from which expenditures are imputed) for each car, ensuring a mechanical connection. Table 8 reports the weighted and unweighted versions of these regressions showing the greater impact of vehicle controls in the NHTS.

Overall, the comparison of the CEX with the NHTS suggests that there are some notable differences in estimated gasoline expenditure, though in most cases there is not a large

difference in predictability within the two samples. While mismeasured expenditures in the CEX may imply that the R^2 is artificially low as compared to some theoretical baseline, it is worth emphasizing a final time that trouble measuring consumption (and hence the burden of a tax) actually makes targeting transfers accurately more difficult.

The Residential Energy Consumption Survey

This paper focuses on gasoline taxes, but it also briefly presents results on home energy consumption. The data quality of the home energy consumption variables in the CEX can be explored by comparison with the Residential Energy Consumption Survey (RECS), which is most recently available in 2009 and 2015. The RECS has the key advantage that electricity and natural gas expenditures are validated against billing records, so the data quality are much better for those variables than in most surveys.

Figure 7 shows the distribution of electricity and natural gas expenditures in the CEX and RECS, pooled for 2009 and 2015. Overall, the similarity in the distributions is broadly encouraging, but there are differences. The CEX shows more observations with low consumption, especially for gas. It also has a longer right tail. This may be in part because the CEX consumer units are on average smaller, but it may also be evidence of mismeasurement.

Figure 7: Comparison of Distribution of Implied Electricity and Natural Gas Expenditures in CEX and RECS

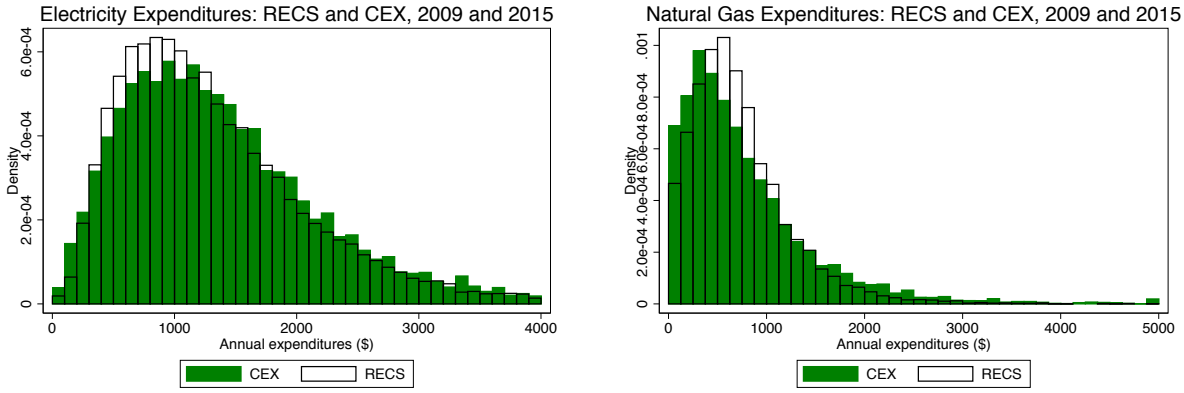

Figure shows histogram of estimated annual fuel expenditure by households using CEX and NHTS data. Left panel is for households with one vehicle. Right panel is for households with two vehicles. Both are from 2009 surveys.

The primary concern with mismeasurement for the core purposes of this study is that it might artificially deflate the degree of predictability. Table 9 shows the R^2 from regressions with the overlapping common set of covariates between the RECS and CEX. The RECS does show a somewhat higher R^2. The data are not winsorized in these regressions. In other specifications (not shown), truncating the right tail of the distribution for high values has little effect on the R^2. Again, the high level point that consumption will be hard to measure and predict is reinforced if the CEX has measurement problems, though it certainly opens the possibility of using better measured surveys to design the transfer system.

Table 9: Predictability of Home Energy Expenditures in CEX versus NHTS

	Electricity				Natural Gas			
	CEX	RECS	CEX	RECS	CEX	RECS	CEX	RECS
R^2	.198	.262	.180	.263	.123	.218	.114	.178
Base controls	Y	Y	Y	Y	Y	Y	Y	Y
Weighted	N	N	Y	Y	N	N	Y	Y
N	17,802	17,769	17,802	17,769	11,263	10,798	11,263	10,798

Table compares 2009 and 2015 CEX to 2009 and 2015 RECS. Dependent variable is annualized expenditures on electricity or natural gas. Base controls include income, Census regions, urban dummy, family size and number of persons over 18. Samples are restricted to households with positive expenditures for natural gas.