Externalities, Internalities, and the Targeting of Energy Policy

Hunt Allcott, Sendhil Mullainathan, and Dmitry Taubinsky*

March 5, 2012

Abstract

We show how the traditional logic of Pigouvian externality taxes changes if consumers are inattentive to energy costs when buying energy-using durables such as cars and air conditioners. First, with inattention, there is an "Internality Dividend" from externality taxes: aside from reducing the provision of public bads, they also reduce allocative inefficiencies caused by consumers' underinvestment in energy efficient durables. Second, although Pigouvian taxes are clearly the preferred policy mechanism when externalities are the only market failure, inattention provides an "Internality Rationale" for alternative policies such as product subsidies that reduce the relative price of energy efficient durables. However, when some consumers misoptimize and others do not, a crucial quantity for policy analysis is the "average marginal internality": the extent to which a policy preferentially targets misoptimizing consumers. We present empirical data suggesting that some energy efficiency subsidies are poorly targeted, as they are primarily taken up by environmentalist consumers that may be less likely to be inattentive to energy costs. As an example of the importance of the average marginal internality, we carry out a randomized field experiment to provide rebates for energy efficient lightbulbs and illustrate how the welfare effects of the rebate depend significantly on whether the more inattentive consumers are more or less elastic. The results show how energy efficiency subsidies and analogous "internality tax" policies, if poorly targeted, can decrease welfare even if the average consumer overconsumes a good and the policy reduces consumption.

JEL Codes: D03, D04, D11, H21, H22, H23, L51, L62, L97, Q41, Q48

Keywords: Energy efficiency, energy-using durables, implied discount rates, inattention, internality taxes, nudges, salience, targeting, average marginal internality.

We thank Meghan Busse, Liran Einav, Larry Goulder, Michael Greenstone, Garth Heutel, Ryan Kellogg, Jim Sallee, Glen Weyl, Nathan Wozny, and seminar participants at the 2011 ASSA meetings, Duke, the Environmental Defense Fund, Harvard, MIT, NBER, New York University, the Stanford Institute for Theoretical Economics, UC Santa Barbara, Universidad de los Andes, and Yale for feedback and helpful conversations. Thanks also to Jermiah Hair, Rich Sweeney, Nina Yang, and Tiffany Yee for excellent research assistance. We are grateful to the Sloan Foundation for financial support of our research on the economics of energy efficiency.

*Allcott: New York University and NBER. Mullainathan: Harvard University and NBER. Taubinsky: Harvard University.

1 Introduction

Since a seminal paper by Hausman (1979), it has frequently been asserted that consumers are inattentive, or otherwise undervalue, energy costs when purchasing energy-using durables such as cars and air conditioners. Although the empirical evidence is under continued debate, this assertion would be consistent with findings that we are inattentive to other ancillary product costs such as sales taxes (Chetty, Looney, and Kroft 2009), shipping and handling charges (Hossein and Morgan 2006), and the out-of-pocket costs of insurance plans (Abaluck and Gruber 2011). Consumer inattention to energy costs has become an important policy issue: it is a key justification for significant regulations such as Corporate Average Fuel Economy (CAFE) Standards and for billions of dollars in subsidies for energy efficient durable goods. Despite the policy importance, however, there is very little formal theoretical guidance on the implications of inattention for the design of energy policy.

In this paper, we take as given the assertion that consumers are partially inattentive to energy costs and ask a set of basic questions. First, how can we formalize inattention and do welfare analysis of policies that are proposed to address this form of misoptimization? Second, how does inattention theoretically change the familiar logic of Pigouvian taxation of energy use externalities, and how large might these changes be in practice? Third, while some consumers may be inattentive, others may not be. What are the welfare consequences of this form of heterogeneity, and how can policy be designed to preferentially target the consumers that misoptimize without distorting decisions of consumers that do not?

Our paper begins with a theoretical analysis of consumers that choose between two energy-using durable goods. One of them, which can be thought of as the "gas sipper," has lower energy costs compared to the other, the "gas guzzler." Consumers have some distribution of utilization demand: some have long commutes, while others live close to the office. When choosing between the two goods, some consumers misoptimize: while they should be indifferent between \$1 in purchase price and \$1 in energy costs because both equally affect consumption of the numeraire good, some consumers undervalue energy costs because these costs are not salient or are paid incrementally in the future. In the language of Herrnstein, Loewenstein, Prelec, and Vaughan (1993), misoptimizing consumers impose an "internality" on themselves: they choose the gas guzzler despite the fact that they would have higher experienced utility with the gas sipper. Consumers also impose externalities on others when they use energy, for example by contributing to climate change and local air pollution. To address these two inefficiencies, the policymaker has two instruments: energy taxes, by which we mean carbon taxes, cap-and-trade programs, gas taxes, and other policies that change the retail energy price, and "product subsidies," by which we mean subsidies for hybrid vehicles, home weatherization, and energy efficient appliances, average fuel economy standards, feebates, and other policies that affect the relative purchase price of the gas sipper vs. the gas guzzler.

We show that adding internalities reverses two results from a canonical Pigouvian framework where energy use externalities are the only market failure. The first canonical result is that while

Pigouvian taxes increase social welfare, they reduce "consumer welfare," by which we mean social welfare with zero weight placed on the externality reduction. In the current context of climate change policy, this traditional result is extremely relevant: some policymakers place little weight on the externality reduction from a carbon tax and argue against such a policy because it damages the economy in the short term. However, we show that in a world with inattention, this result is reversed: a carbon tax can actually increase consumer welfare, independent of the reduction in externalities. Intuitively, this is because inattention is a pre-existing distortion that increases demand for gas guzzlers above consumers' private optima, and increasing energy taxes helps to correct this distortion. Conceptually, this result is related to the Double Dividend hypothesis explored by Bovenberg and Goulder (1996) and many others, in the very basic sense that it identifies an additional benefit from environmental taxation other than externality reduction. As such, we call this the effect the *Internality Dividend from Energy Taxes.*

The second canonical result from the Pigouvian framework is that when energy use externalities are the only market failure, product subsidies are an inefficient second best substitute for Pigouvian energy taxes (Jacobsen 2010, Krupnick *et al.* 2010, Sallee 2011a). One key reason for this is that unlike energy taxes, product subsidies do not impose the correct social cost of energy use on consumers' utilization decisions: while product subsidies can induce consumers to buy the first best quantity of gas sippers, consumers will still drive too much. However, we show that inattention to energy costs is more effectively addressed with product subsidies than with higher energy taxes. Intuitively, one reason for this is that while the energy tax can indeed be raised high enough to induce partially attentive consumers to purchase the first best quantity of gas sippers, once they own the vehicles they drive too little because the tax-inclusive energy price is too high. Thus, once there are both externalities and internalities, these two sources of inefficiency are best addressed through two instruments: the energy tax targets the externality, while the product subsidy targets the internality. We call this the *Internality Rationale for Product Subsidies.* Because product subsidies are effectively "internality taxes" on gas guzzlers, this result parallels the O'Donoghue and Rabin (2006) finding that a policymaker would optimally impose sin taxes if consumers misoptimize by over-consuming a sin good.

In order to provide formal theoretical guidance on how thes several billion dollars of energy efficiency subsidies being disbursed each year in the U.S. might be set optimally, we derive a formula for the optimal product subsidy under inattention. Our result is quite intuitive: the optimal internality tax equals the weighted average of the internalities of the marginal consumers, with weights related to the price derivatives of demand. This is analogous but not identical to the Diamond (1973) result that the optimal externality tax in the presence of heterogeneous externalities is the weighted average of the externalities, the weights being the price derivatives of demand.

To complement the theoretical analysis, we calibrate a discrete choice model of U.S. automobile demand with inattentive consumers, using a set of utility function parameters from the literature. We calibrate the Internality Dividend from Energy Taxes, showing that increasing the energy tax by the estimated climate change externality *increases* consumer welfare by \$6.70 per ton of carbon dioxide abated. We also document the Internality Rationale for Product Subsidies by calibrating the socially optimal combination of product subsidies and energy taxes. The optimal product subsidy in that case is significant: it increases the relative price of a 20 mile-per-gallon (MPG) vehicle such as the Subaru Outback by \$800 relative to a 25 MPG vehicle such as the Toyota Corolla.

So far, these results are qualitatively consistent with existing policy arguments that inattention justifies energy efficiency subsidies. Imagine the very plausible scenario in which some consumers are inattentive, while others are not. An increasingly large product subsidy will increasingly distort decisions by rational consumers, even as it generates allocative gains for inattentive consumers. On net, the effect of a marginal increase in a policy on consumer welfare is the average of the internalities of consumers *whose choices are changed*. Thus, it is not the average population internality that matters: it is the *average marginal internality*. Even in a world where the average consumer is highly inattentive to energy costs, subsidizing energy efficient durables can be undesirable if the subsidy primarily affects choices for attentive types. Ideally, a policy would be well-targeted: it would preferentially change decisions of misoptimizing consumers without distorting rational consumers. This concept of behavioral targeting is a microfounded version of the logic behind "nudging" (Thaler and Sunstein 2008), libertarian paternalism (Sunstein and Thaler 2003), and asymmetric paternalism (Camerer *et al.* 2003), which entail policies that affect misoptimizers without affecting rationals.

This motivates us to examine data on whether existing energy efficiency policies are well-targeted. Specifically, we document the observable characteristics of consumers who take up subsidies for energy efficient appliances, and ask whether these observable characteristics are likely to be correlated with inattention. We show that subsidy beneficiaries are more likely to have installed solar panels on their homes and voluntarily pay for green energy. These data echo Kahn's (2007) empirical finding that hybrid vehicles, which benefit from subsidies and are an important way that automakers comply with CAFE standards, are preferentially driven by consumers in "green" zip codes. In effect, it appears that instead of changing decisions by consumers that appear to be inattentive to energy costs, energy efficiency subsidies encourage already-green consumers to be even more green. Multiple mechanisms could explain this poor targeting: for example, inattentive consumers may often be unaware energy efficiency subsidies, or they could rationally have weaker preferences for green products and thus would tend to be farther from the margin.

Finally, we illustrate the importance of targeting in the context of a field experiment in which we provide rebates for energy efficient Compact Fluorescent Lightbulbs (CFLs) to randomly-selected consumers at a major retailer. We characterize consumers as either less or more attentive based on whether they report that energy costs are an important factor in their lightbulb purchase decision. We use the empirical results to generate upper and lower bounds on the differential elasticities of less-attentive versus more-attentive consumers. We show that the optimal subsidy and the welfare effects thereof significantly on whether the rebate has greater effect on decisions by more or less attentive types. This example calibration highlights the following basic insight: when consumers are heterogeneous, the welfare analysis depends not just on *how much* energy is saved, but also on *who* is saving the energy.

Although our analysis centers on energy policy, this should not obscure the general importance of behavioral targeting and the average marginal internality. A number of proposed and actual policies, such as taxes on cigarettes, alcohol, and the fat content of foods, are partially justified by the idea that they will benefit individuals who are misoptimizing. In the very plausible world where some consumers misoptimize and others do not, it is crucial to understand the tax elasticities of different consumer types. If misoptimizing types are inelastic and rational types are not, then *the welfare effects of an internality tax could be negative even if the average consumer overconsumes and the tax reduces consumption.* For example, this might be the case when the consumers believed to be misoptimizing are addicted to the sin good, as discussed in Bernheim and Rangel (2005). Fundamentally, our argument is a behavioral economics analogue to the tax targeting literature in other areas of public finance, such as Akerlof (1978).

The paper proceeds as follows. In Section 2, we provide more background on inattention to energy costs and relevant energy efficiency policies. Section 3 presents our theoretical model and formal results on optimal tax policy. Section 4 details the auto market simulation and results. Section 5 discusses the average marginal internality and presents empirical data on the targeting of existing energy efficiency policies. Section 6 presents the lightbulb rebate field experiment and welfare calculations. Section 7 concludes with a discussion of policies that might preferentially target inattentive consumers.

2 Background

2.1 Inattention to Energy Costs

In this paper, we use the word "inattention" to describe why consumers might not choose more energy efficient goods that could make them better off. This is largely semantic: we show in Appendix [tk-forthcoming] that our formal results hold under a broader class of models under which consumer demand for energy efficiency is lower than the level that minimizes private cost. Instead of inattention, we could also potentially use words such as "undervaluation," "myopia," naive hyperbolic discounting," "systematically biased beliefs," "rational inattention," "costly information acquisition," and "shrouded attributes." Here we discuss the empirical literature that has jointly tested such models.

Hausman (1979) estimated that the "implied discount rate" that rationalizes consumers' trade-offs between purchase prices and future energy costs for air conditioners was 15 to 25 percent, above the rates at which most consumers borrowed and invested money. His results were corroborated by Gately (1980), who showed that buyers of energy inefficient refrigerators needed to have discount rates of 45% to 300%, and by Dubin and McFadden (1984), who found that choices and utilization of home heating equipment implied a 20 percent discount rate. Hausman (1979) argued that consumers were making mistakes by not buying more energy efficient appliances, but that this was unsurprising because "at least since Pigou, many economists have commented on a 'defective telescopic faculty.'"

A number of papers have tested for whether automobile consumers appear to undervalue future gasoline costs relative to purchase prices, including Austin (2008), Dreyfus and Viscusi (1995), Espey and Nair (2005), Goldberg (1998), Kilian and Sims (2006), Sawhill (2008), and Verboven (1999, 2002). Greene (2010) reviews 25 studies, of which 12 suggest that consumers undervalue gas costs, five suggest that we overvalue gas costs, and eight indicate that the average consumer makes the tradeoff correctly. The most recent estimates in this literature, by Allcott and Wozny (2011), Busse, Knittel, and Zettelmeyer (2011), and Sallee, West, and Fan (2011), identify off of changes in gasoline prices over time, which is appealing because it allows them to relax the commonly-required assumption that energy efficient models do not have unobserved attributes that make them more or less appealing to consumers. These recent papers highlight the difficulty in testing empirically whether the average consumer misoptimizes: Allcott and Wozny's base specification suggests that consumers somewhat undervalue gasoline costs, but they show how different plausible assumptions can affect their results. Busse, Knittel, and Zettelmeyer (2011) present a set of results that could suggest that consumers undervalue, correctly value, or overvalue gasoline costs. They argue that taken as a whole, their results provide little evidence that the average consumer undervalues.

The idea that consumers are inattentive to a product's energy costs would be consistent with empirical evidence from other domains that we are inattentive to other ancillary product costs. Consumers on eBay, for example, are less elastic to shipping and handling charges than to the listed purchase price (Hossain and Morgan 2006). Mutual fund investors appear to be less attentive to ongoing management fees than to upfront payments (Barber, Odean, and Zheng 2005). Chetty, Looney, and Kroft (2009) show that shoppers are less elastic to sales taxes than to prices. Abaluck and Gruber (2011) show the seniors choosing between Medicare Part D plans place more weight on plan premiums than on expected out-of-pocket costs.

2.2 Existing Energy Taxes, Subsidies, and Standards

In the U.S., a wide array of state and federal policies encourage energy efficiency. Our analysis focuses specifically on what we call "product subsidies": taxes or subsidies that reduce the relative prices of energy efficient durable goods. Such policies include tax credits of up to $3400 for hybrid vehicles, which were available for the bulk of the last decade, as well as the "gas guzzler tax," an excise tax ranging from $1000 to $7700 on the sale of low fuel economy passenger cars. Another example is the Weatherization Assistance Program, which heavily subsidizes weatherization for about 100,000 low-income homeowners each year. Furthermore, in many states, there are an array of rebates and subsidized loans for weatherization and energy efficient appliances; these "Demand-Side Management programs" cost about $3.6 billion per year (U.S. EIA 2010).

Importantly, our model of product subsidies also captures the effects of the Corporate Average Fuel Economy (CAFE) standard. This policy requires that the fleets of new cars and trucks sold by each auto manufacturer attain a minimum average fuel economy rating. This constraint adds a relative shadow cost to the sale of low fuel economy vehicles, inducing automakers to increase their relative prices. Thus, the CAFE standard affects consumers in the same way as a product subsidy, by changing relative product purchase prices. In the long run, of course, both explicit subsidies and the CAFE standard induce changes in the characteristics of vehicles offered, but this is well beyond the scope of our analysis. Our study builds on other studies of CAFE standards and other potential policies to decrease the relative purchase prices of energy efficient vehicles, including Anderson, Parry, Sallee, and Fischer (2010), Austin and Dinan (2005), Fischer, Harrington, and Parry (2007), Fullerton and West (2010), Goldberg (1998), Greene, Patterson, Singh, and Li (2005), Jacobsen (2010), Kleit (2004), Sallee (2011a), and Davis, Levine, and Train (1993).

Since the mid-1980s, there have also been a set of minimum energy efficiency standards for refrigerators, air conditioners, hot water heaters, and many other household appliances. These standards can also be captured in our model, as product subsidies so large as to completely eliminate sales of the energy inefficient good. Since the set of policy options we consider places no restrictions on the level of the product subsidy, we know that a minimum energy efficiency standard is not the optimal policy if the optimal product subsidy is below the level that implies zero sales of the energy inefficient good. However, our model does not include a cost of public funds, which should factor into a full evaluation of product taxes or subsidies versus minimum energy efficiency standards.

Why do we have these policies? As discussed in Allcott and Greenstone (2012), potential reasons include externalities, internalities, and a set of other market failures largely deriving from imperfect information that could cause consumers and firms to underinvest in energy efficient goods. The informal policy logic is well-summarized in Hausman's (1979) discussion of consumers' high estimated discount rate: "Since this individual discount rate substantially exceeds the social discount rate used in benefit-cost calculations, the divergence might be narrowed by policies which lead to purchases of more energy-efficient equipment." The idea that energy efficiency policies can correct consumer misotimization plays an important role in some discussions, including a central role in the U.S. government's official cost-benefit analysis of recent increases in the CAFE standard.[1] Several other analysis explore this idea theoretically or analytically, including Allcott and Wozny (2011),

[1] In its Regulatory Impact Analysis of the recently strengthened CAFE standard (2010, page 2), the National Highway Traffic Safety Administration (NHTSA) writes, "Although the economy-wide or "social" benefits from requiring higher fuel economy represent an important share of the total economic benefits from raising CAFE standards, NHTSA estimates that benefits *to vehicle buyers themselves* [original emphasis] will significantly exceed the costs of complying with the stricter fuel economy standards this rule establishes . . . However, this raises the question of why current purchasing patterns do not result in higher average fuel economy, and why stricter fuel efficiency standards should be necessary to achieve that goal. To address this issue, the analysis examines possible explanations for this apparent paradox, including discrepancies between the consumers' perceptions of the value of fuel savings and those calculated by the agency . . . "

Fischer, Harrington, and Parry (2007), Krupnick *et al.* (2010), and Parry, Evans, and Oates (2010). In particular, Heutel (2011) is a nice related paper that studies command-and-control vs. market based environmental regulation under hyperbolic discounting; our Proposition 5 is analogous to his Propositions 2 and 3.

Different readers will have different assessments of the empirical evidence on whether consumers misoptimize, as well as different philosophies on whether it is even possible for individuals to make a mistake and whether policymakers should intervene if they do. However, the fact is that energy policies that cost many billions of dollars are partially or even largely justified as responses to some form of consumer misoptimization. This paper is motivated by the idea that aside from empirically testing if and when consumers misoptimize, it is also crucially important to provide formal theoretical analysis that can help improve the design of these policies. We begin that in the next section.

3 Optimal Taxation of Energy-Using Durables

3.1 Setup

3.1.1 Consumer Utility

We model consumers who choose between an energy inefficient durable I, and an energy efficient durable E. Concretely, we have in mind a choice between hybrid versus non-hybrid cars, compact fluorescent lightbulbs versus incandescents, and standard versus energy efficient versions of air conditioners, washing machines, and other appliances. Consumers have single unit demand, and the durables differ in their energy efficiency. A durable $j \in \{I, E\}$ consumes e_j units of energy per unit of utilization m, with $e_I > e_E$.

Consumers are differentiated by a parameter θ, which corresponds to how much a consumer will utilize his durable. A high-θ consumer is one who has a long commute to the office or lives in a hot climate that requires lots of air conditioner use. We assume that each consumer chooses a utilization level $m > \theta$, from which he derives utility $u(m - \theta)$. To ensure the existence of an interior optimum, we assume $u' > 0$, $u'' < 0$, $\lim_{x \to 0} u'(x) = -\infty$ and $\lim_{x \to 0} u'(x) = 0$. We also assume that $|xu''(x)/u'(x)| > 1$ to ensure that the price elasticity of energy use is less than 1 for each consumer and that consumers use less energy when they purchase the more energy efficient durable. The parameter θ is distributed according to some atomless distribution F with positive support on the positive reals.

For simplicity, we assume that the two durable goods differ only in energy efficiency and not in how they directly impact a consumer's utility. We also assume that there is no outside option. We abstract away from the outside option for two reasons. First, this allows us to remain agnostic about how exactly consumer inattentiveness to differences in energy costs impacts their choice of an outside option. Second, this also allows us to interpret our model as a model of consumer choice of efficiency enhancements such as weatherization. Indeed, I can be viewed as the status quo of all consumers who have not weatherized their homes, whereas E is the improved efficiency of consumers who have weatherized their homes.

Whatever consumers don't spend on purchasing the durable and subsequent energy use, they spend on the numeraire good. Therefore, if p_g is the cost of energy, p_j is the price of durable j, T is a transfer from the government and Y is the budget constraint, then a consumer derives utility

$$\{Y + T - p_j - p_g m e_j\} + u(m - \theta) \tag{1}$$

from purchasing durable j and choosing m units of utilization. Notice that the term in brackets is consumption of the numeraire good: the amount of money from income Y and transfers T that the consumer has left over after purchasing the durable good and paying for energy. Each consumer's budget constraint is large enough so that the optimal choice m^* is an interior solution.

3.1.2 Consumer Choice

We assume that while a consumer's utility is determined by θ alone, consumer choice may also be driven by a "mis-optimization" or attention parameter γ.

It is helpful to define the function v as follows:

$$v(\theta, e, p_g) \equiv \max_m \{u(m - \theta) - p_g m e\}. \tag{2}$$

Think of $v(\theta, e_E, p_g) - v(\theta, e_I, p_g)$ as the gross utility gain from the energy efficient good, and $p_E - p_I$ as the incremental price. A fully optimizing consumer chooses durable E if and only if

$$v(\theta, e_E, p_g) - v(\theta, e_I, p_g) > p_E - p_I. \tag{3}$$

Mis-optimizing consumers, on the other hand, are not fully attentive to how differences in energy efficiency will impact their future utility, and choose E if and only if

$$\gamma[v(\theta, e_E, p_g) - v(\theta, e_I, p_g)] > p_E - p_I \tag{4}$$

for some $\gamma \in (0, 1)$.

The inattentive consumers in our model are similar to "myopic" consumers in Gabaix and Laibson (2006) that do not fully value "add-on costs" when purchasing a good or service. These consumers do not rationally acquire information about add-on costs or rationally infer their magnitude. However, as we show in Appendix [forthcoming-tk], our basic results go through in a broader class of models, as long as there is initial underinvestment in energy efficiency.

We will use the following additional notation throughout the paper: p will refer to the price vector (p_I, p_E, p_g) and $\xi(\theta, \gamma, p)$ will denote the consumer's choice of durable I or E (at prices p).

3.1.3 The Government

Products $j \in \{E, I\}$ are produced in a competitive economy at a constant marginal cost c_j, with $c_I < c_E$. Similarly, energy is produced in a competitive market at constant marginal cost c_g. The government chooses a subsidy τ_E for product E and an energy tax τ_g.[2] Prices are then given by $p_I = c_I$, $p_E = c_E - \tau_E$, $p_g = c_g + \tau_g$. We will use τ to refer to the tax policy vector (τ_E, τ_g), and use $T(\tau)$ to refer to the tax revenue from that policy (which could be negative).

The government maintains a balanced budget. Because $T(\tau)$ is a lump-sum tax or transfer, taxing or subsidizing durables purchases or energy use has no distortionary effects on other dimensions of consumption. We are therefore abstracting to a simplified scenario in which the cost of public funds is 1.

Define ϕ as the marginal damage per unit of energy used, $Q_g(p)$ as the amount of energy used at prices p, and H as the joint distribution of (θ, γ). For a consumer of type (θ, γ), also define

$$V(j, \theta, \gamma) \equiv v(\theta, e_j, p_g) - p_j \tag{5}$$

to be the experienced utility from purchasing durable j. Notice that for $\gamma \neq 1$, consumers misoptimize and therefore do not necessarily choose j to maximize $V(j, \theta, \gamma)$. The government wishes to set τ so as to maximize consumer utility net of the damage caused by energy use:

$$W(\tau) \equiv \int [V(\xi(\theta, \gamma, p), \theta, \gamma) + Y + T(\tau)] dH - \phi Q_g(p). \tag{6}$$

We will call W the *social welfare* and call $W^{SB} \equiv \max_\tau W(\tau)$ the *second best*. We will use W^{FB} to refer to the *first best*—the maximum social welfare that is obtainable under any possible combination of choices of durables and utilizations by consumers.[3]

At times we will be interested in a slightly different objective function that doesn't consider the marginal damage and focuses solely on consumer utility. We use W_0 to denote this objective function and define it exactly the same way as W except without the final term $\phi Q(p)$. We will refer to W_0 as *consumer welfare*. Unless otherwise stated, however, we focus our analysis on the social welfare W.

Figure 1 illustrates the setup of equilibrium in the durable goods market. The two goods are supplied perfectly elastically, and the incremental price of good E is the horizontal black line. The first best demand curve, if consumers all have $\gamma = 1$, is the solid blue line through points c and a. The shape of the demand curve is determined by the distribution of gross utility gain from good E, $v(\theta, e_E, p_g) - v(\theta, e_I, p_g)$, which itself is determined by the distribution of utilization needs θ. The first best equilibrium is at point a, with quantity demanded q^*. For the marginal consumer at that point, the gross utility gain just equals the incremental price. However, inattentive consumers

[2] We do not lose any generality by not considering only the subsidy for E. In our model, subsidies τ'_I and τ'_E for products I and E, respectively, are choice and welfare equivalent to subsidies $\tau_I = 0$, $\tau_E = \tau'_E - \tau'_I$.

[3] To be more precise, set $w(\theta) \equiv \max_{m, i \in \{I,E\}} \{u(m - \theta) - (c_g + \phi)m - c_i\}$. Then $W^{FB} = \int w(\theta) dF$.

undervalue the gross relative utility gain $v(\theta, e_E, p_g) - v(\theta, e_I, p_g)$ by factor $\gamma < 1$, and their demand curve for good E shifts downward proportionally. The equilibrium under inattention is at point b, and the consumer welfare loss from inattention is the triangle abc.

3.2 The Internality Dividend from Energy Taxes

To keep our results simple and sharp, we work with a simple distribution of attention in which a fraction α of consumers have attention parameter $\gamma_L \in (0, 1]$ and a fraction $(1-\alpha)$ of consumers have attention parameter $\gamma_H \in [\gamma_L, 1]$. The distribution of attention is independent of the distribution of θ.

A canonical result is that when consumers perfectly optimize, their welfare W_0 (which does not take into account damages from energy use) cannot be increased with taxes, since in our framework these can only be distortionary. Similarly, when consumers optimize perfectly, social welfare W (which takes damages into account) is maximized simply by setting equating the energy tax to the marginal damage. We note this as Claim 1:

Claim 1 *Suppose that consumers optimize perfectly ($\gamma_L = \gamma_H = 1$). Then then consumer welfare W_0 is maximied by $\tau_g^* = 0$ and $\tau_E^* = 0$. Similarly, social welfare W is maximized by $\tau_g^* = \phi$ and $\tau_E^* = 0$.*

Notice how in the model with externalities only, the Pigouvian tax $\tau_g^* = \phi$ increases social welfare but reduces consumer welfare.

When there are some inattentive consumers, however, some additional intervention is optimal even when energy use externalities are not taken into account. When at least some consumers underconsume E, it is optimal to encourage more purchase of E with either a subsidy or a higher energy tax. In particular, if the government does not rely on subsidies, then a higher energy tax improves consumer welfare.

Proposition 1 *Suppose that $\gamma_L < 1$. If the government maximizes W_0 then the energy tax that maximizes consumer welfare is $\tau_g^{**} > 0$.*

The basic intuition behind this proposition is that inattention is a pre-existing distortion that reduces demand for the energy efficient good E below consumers' private optima. A positive energy tax induces some consumers that had misoptimized by choosing good I to instead choose good E, increasing consumer welfare. It should be emphasized that this proposition holds even if $\gamma_H = 1$. That is, even if some consumers choose optimally, then additional intervention is still beneficial, even at the cost of making these consumers' choices less efficient. The reason is that if a consumer with attention γ_H is indifferent between E and I at the policy $(\tau_E, \tau_g) = (0, 0)$, then the benefit of giving E to this consumer equals the benefit of giving I to this consumer. Thus the efficiency loss from changing the choices of optimizing consumers who are close to indifferent between E and

I is first-order zero. On the other hand, the gain to encouraging more consumers with $\gamma_L < 1$ to purchase E is first-order positive. This intuition, which is similar to the basic logic underlying the Envelope Theorem, is emphasized by O'Donoghue and Rabin (2006) in their analysis of optimal sin taxes.

This proposition illustrates how inattention reverses the traditional result that energy taxes reduce consumer welfare. Why is this important? One reason is that some policymakers argue against carbon taxes or other energy taxes because they are "bad for the economy," which in our model formally means that they are bad for consumer welfare. Our result shows that even a policymaker who places zero importance on externality reductions might still support an energy tax. This result relates to the Double Dividend hypothesis in the basic sense that it identifies a potential benefit of environmental taxation other than externality reduction. As such, we call this the *Internality Dividend from Energy Taxes.*

Figure 2 illustrates how an energy tax increases consumer welfare. The setup is the same as in Figure 1. For simplicity, imagine that all consumers have homogeneous $\gamma < 1$ such that the dashed red line is the market demand curve, and q_L is the quantity demanded of E. An energy tax rotates up the demand curve, shifting the equilibrium to point d. The set of consumers between q_L and q'_L now purchase good E, as they do in the first best, and consumer welfare is higher. Although consumers also pay more in taxes, this money is recycled to them through transfer T. The energy tax that maximizes consumer welfare trades off these gains from improved product allocation with the allocative losses from reduced utilization due to higher energy prices.

3.3 The Internality Rationale for Product Subsidies

Claim 1 reminds us that when externalities are the only market failure, not only does the optimal energy tax give the first best, but the optimal product subsidy is zero. In practice, the evidence suggests that product subsidies are a highly inefficient substitute for the energy tax. For example, Krupnick *et al.* (2010) show that proposed energy efficiency standards have five times more consumer welfare cost per ton of carbon abated than energy taxes, and Jacobsen (2010) shows that CAFE standards cost 2.5 times more per ton abated than gas taxes. One of the main reasons for this is that if the energy price is not at the first best level, a product subsidy will cause consumers to buy more energy efficient goods but then use them too much.

In this section, however, we show that when consumers are inattentive, product subsidies can now increase welfare. Furthermore, the policy problem is essentially one of two market failures and two instruments, where the energy tax primarily targets the externality and the product subsidy primarily targets the internality. We call this logic the *Internality Rationale for Product Subsidies.*

The basic Internality Rationale result is that under inattention, an optimal combination of subsidy and energy tax must include a positive product subsidy:

Proposition 2 *Suppose that* $\gamma_L < 1$. *Then* $\frac{\partial}{\partial \tau_E} W > 0$ *and* $\frac{\partial}{\partial \tau_g} W > 0$ *at* $(\tau_E, \tau_g) = (0, \phi)$. *If*

(τ_E^*, τ_g^*) *is an optimal tax policy, then* $\tau_E^* > 0$.

We now show that under certain conditions, more inattention implies that product subsidies are more "important" in two senses. First, we show that holding heterogeneity constant, more inattention implies a larger product subsidy.[4] More formally, consider two different distributions of γ, G and G', that have attention weights $\{\gamma_H, \gamma_L\}$ and $\{\gamma'_H, \gamma'_L\}$, respectively. G' implies more inattention: $\gamma'_L < \gamma_L$. Suppose that "attention heterogeneity" is the same in these two distributions: α is the same, and $\gamma_H/\gamma_L = \gamma'_H/\gamma'_L$. Proposition 3 shows that the optimal product subsidy is larger under G':

Proposition 3 *Suppose that* (τ_E^*, τ_g^*) *is an optimal tax policy under G, and suppose that* $\tau_E^{**} > \tau_E^*$ *satisfies* $c_E - c_I - \tau_E^{**} = \frac{\gamma'_L}{\gamma_L}(c_E - c_I - \tau_E^*)$. *Then* (τ_E^{**}, τ_g^*) *is an optimal tax policy under* G'.

The second sense in which increasing inattention makes the product subsidy more important is that the social welfare that can be achieved by the energy tax alone is decreasing in the inattentiveness of the consumers. Define W_{energy}^{TB} to be the "third-best" level of social welfare that can be achieved by the energy tax alone when the subsidy is constrained $\tau_E = 0$.

Proposition 4 W_{energy}^{TB} *is smaller under* G' *than under* G *and* $W^{SB} - W_{energy}^{TB}$ *is larger under* G' *than under* G.

Essentially, these two sources of inefficiency require two corrective instruments. Why is the energy tax alone ineffective at addressing internalities? One key reason is the utilization elasticity. If utilization demand were fully inelastic, an energy tax could be set higher and higher to correct for increasing inattention, as long as consumers are not fully inattentive to energy costs. The problem with this approach is that the increasingly large energy tax increasingly distorts utilization choices away from the first best: consumers buy more energy efficient goods but use them too little. Thus, the problem with using energy taxes as an instrument to address inattention to energy costs is not just that consumers are inattentive to the tax: it is that the energy tax also distorts decisions on a second margin where consumers are assumed to be fully attentive.

Given that internalities provide a rationale for some product subsidy, what is the optimal product subsidy? We now derive a formula for the optimal subsidy given any energy tax τ_g. Two pieces of notation are required. First, let D_i denote the total demand for product E by consumers with attention weight γ_i, and let D'_i be the derivative of D_i with respect to τ_E. Second, let $F_i D'_i$ be the marginal change in total energy consumed when the subsidy τ_E is perturbed.

[4] We emphasize the importance of holding heterogeneity constant, as the simple intuition that more inattention calls for more intervention is not necessarily correct. Consider, for example, the effect of varying γ_L while γ_H is fixed at $\gamma_H = 1$. For intermediate values of γ_L, the optimal intervention might be quite sizable. However, as γ_L gets close to zero so that the less attentive consumers are nearly insensitive to the advantages of purchasing E, any taxes that fall short of making $p_I \approx p_E$ will have very little effect on the less attentive consumers. To make this effect very clear, consider the limit case $\gamma_L = 0$, so that unless $p_I = p_E$, consumers will not purchase E. Thus any intervention that impacts the choices of the γ_L consumers forces all consumers with $\gamma_H = 1$ to purchase E. So if there are enough consumers with $\gamma_H = 1$, then no intervention may be optimal at all.

By Equation (4), the social benefit of obtaining E rather than I to the γ_i consumer who thinks he is indifferent between E and I is $(c_E - c_I - \tau_E)/\gamma_i - (c_E - c_I) + F_i(\tau_g - \phi)$. Thus the total impact of a marginal increase in τ_E is given by

$$\sum_i \left[(c_E - c_I - \tau_E)/\gamma_i - (c_E - c_I) + (\tau_g - \phi)F_i\right] D_i' \tag{7}$$

The optimal value of τ_E is determined by setting the expression in 7 equal to zero:

$$\tau_E = (c_E - c_I)\left(1 - \frac{\sum_i D_i'}{\sum_i \frac{D_i'}{\gamma_i}}\right) + (\phi - \tau_g)\frac{\sum_i F_i D_i'}{\sum_i \frac{D_i'}{\gamma_i}}. \tag{8}$$

The above equation covers the case when the energy tax does not equal marginal damages. When $\phi = \tau_g$, we obtain a simpler expression:

$$\tau_E = (c_E - c_I)\left(1 - \frac{\sum_i D_i'}{\sum_i \frac{D_i'}{\gamma_i}}\right) \tag{9}$$

To build additional intuition, consider the case when consumers have homogeneous inattention, i.e. $\gamma_L = \gamma_H \equiv \gamma$. Furthermore, recall that there is an optimal marginal utilization type θ^* for whom the gross utility gain from the energy efficient product is just equal to the incremental cost: $v(\theta^*, e_E, p_g) - v(\theta^*, e_I, p_g) = (c_E - c_I)$. This means that the optimal product subsidies under homogeneous γ is:

$$\tau_E = (v(\theta^*, e_E, p_g) - v(\theta^*, e_I, p_g))(1 - \gamma) \tag{10}$$

This shows the intuitive result that the optimal product subsidy is equal to the internality of the consumer who is marginal in the social optimum. It corrects relative product prices by exactly the amount that this optimal marginal consumer misperceives the gross utility gains.

One can also see that Equation (9), the optimal uniform internality tax in the presence of heterogeneous internalities is the weighted average of internalities, the weights being related to the price derivatives of demand. Although not identical, this is roughly analogous to the result of Diamond (1973), who shows that the optimal uniform externality tax in the presence of heterogeneous externalities is the weighted average of the externalities, the weights being the price derivatives of demand.

Notice that the optimal value of τ_E is positive as long as $\gamma_i \leq 1$, $\tau_g \leq \phi$, and $\gamma_i < 1$ for at least one i. Notice also how the formula makes explicit that what drives subsidies to be high is not just lower γ_i, but also the share of the different attention types on the margin; that is, $D_i/(\sum_i D_i')$ determines how responsive the optimal subsidy is to the types with attention γ_i. Put differently, what matters here is not the average internality, but an *average marginal internality.*

3.4 The Welfare Effects of Heterogeneity

We now examine the welfare effects of heterogeneity in γ. How close to the first best can we get with tax and subsidy policies that are uniform across consumers, when the internality is non-uniform?

Proposition 5 states that when consumers are homogeneous in their inattention ($\gamma_H = \gamma_L$), a proper choice of subsidy recovers the first best.

Proposition 5 *Suppose that $\gamma_L = \gamma_H \equiv \gamma < 1$. Then the first best is uniquely achieved with $\tau_g^* = \phi$ and $\tau_E^* < 0$. Moreover, the size of the optimal subsidy, $|\tau_E^*|$, is strictly decreasing in γ.*

The basic intuition for the previous proposition can be illustrated by returning to Figure 2. Here again, the line connecting points c and a would be the demand curve if $\gamma = 1$, and the dashed line through point b is the true demand curve for a population of consumers who all have $\gamma_L = \gamma_H = \gamma < 1$. At $\tau_g = \phi$, consumers will choose in a socially efficient way on the intensive margin. However, when $\tau_E = 0$, consumers will underpurchase E relative to the social optimum: the equilibrium quantity will be $q_L < q^*$. A subsidy that reduces the relative price of E to the point where the equilibrium quantity demanded is q^* achieves the first best.

When consumers are heterogeneous in their degree of attention ($\gamma_L \neq \gamma_H$), the first best is no longer possible. Figure 3 illustrates this point. Imagine that the solid blue line is the demand curve for a perfectly attentive subset of consumers with $\gamma = \gamma_H = 1$, and the dashed red line is the demand curve for the subset of inattentive consumers with $\gamma = \gamma_L < 1$. The first best quantity demanded of the energy efficient good is q_H. A subsidy that brings the relative price of E to the dotted horizontal line will improve allocations for inattentive consumers, increasing quantity demanded from q_L to q'_L. However, the subsidy also distorts the decisions of the perfectly attentive types, increasing quantity demanded from q_H to q'_H. The subsidy level drawn in Figure 3 is too large for some consumers and not strong enough for others: there is remaining welfare loss in the white triangle (agh) and the shaded red triangle (amn) relative to the first best. The simple intuition is that a homogeneous subsidy cannot correct misoptimization by heterogeneous types.

Notice that whether or not the first best can be achieved does not depend on how inattentive the agents are, but rather on whether or not they are homogeneous in their inattention. On Figure 3, imagine shrinking the difference in slopes between the two lines. The white and red welfare loss triangles (agh and amn) shrink, and as the heterogeneity disappears, we return to the first best level of Figure 2.

We now show this result formally, using two different ways of thinking about heterogeneity. First, we can ask what happens as we increase or decrease the fraction of less attentive agents in the population. As would be suggested by Proposition 5, when $\alpha \approx 0$ or $\alpha \approx 1$, so that the agents are concentrated around one particular level of attention, the second best should be very close to the first best. As we increase heterogeneity by moving α further away from 1 or from 0, however, the gap between the first and second best increases. This is part 1 of Proposition 6.

Second, we can ask what happens when we broaden the support of the distribution of attention. It turns out that what determines the second best is not the absolute difference $\gamma_H - \gamma_L$ between the highest and smallest levels of attention, but rather the ratio γ_H/γ_L. For example, if $\gamma_L = 0.8$ and $\gamma_H = 0.9$, so that $\gamma_H - \gamma_L = 0.1$ and $\gamma_H/\gamma_L = 1.125$, then the second best may be quite close to the first best. On the other hand, if $\gamma_L = 0.2$ and $\gamma_H = 0.1$, so that $\gamma_H/\gamma_L = 2$, the second best is now much further from the first best, even though we still have $\gamma_H - \gamma_L = 0.1$. Intuitively, this is because the relation between the marginal attentive consumer and the marginal inattentive consumer is determined by γ_H/γ_L. For example, if the marginal attentive consumer assigns twice as much weight to energy costs than the marginal inattentive consumer, then his energy cost savings from purchasing E will be approximately 50% of the energy cost savings of the marginal inattentive consumer. Part 2 of Proposition 6 is that the allocation under the optimal policy is less socially efficient the bigger the difference between the marginal consumers from the different attention groups.

Proposition 6 *Let W^{FB} denote the first best welfare and let W^{SB} be the maximum achievable welfare using taxes τ_E and τ_g. Then*

1. *Holding γ_L and γ_H constant, there is $\alpha^\dagger \in (0,1)$ such that $W^{FB} - W^{SB}$ is increasing in α when $\alpha > \alpha^\dagger$ but decreasing in α when $\alpha < \alpha^\dagger$.*

2. *Holding α constant, $W^{FB} - W^{SB}$ is continuous and strictly increasing in γ_H/γ_L.*

Propositions 6 illustrates one of our main points about heterogeneity and the efficacy of taxes: as consumers become more and more heterogeneous in their levels of attention, tax policy becomes more and more of a blunt instrument. Intuitively, this is because as the distance between different consumers' levels of attention grows, any "compromise" tax policy becomes further from each type's own optimal level.

Heterogeneity also implies that the *targeting* of a policy is important. This is true not just for tax policies, but also information disclosure or any other mechanism in general. To see this mathematically, consider some policy instrument, denoted n. Denote by $D_L(n)$ and $D_H(n)$ the demand curves of the two attention types as a function of n. The social benefit of a marginal increase in the strength of the policy is

$$D_L'(n)b_L + D_H'(n)b_H \tag{11}$$

where b_L and b_H are the marginal social benefits corresponding to the marginal consumer of type L or H purchasing E. Again, we see here that what matters is the average marginal benefit - not the average benefit in the population - of moving consumers to the efficient good. As illustrated by Figure 3, $b_L > b_H$: the marginal low attention type is making a larger mistake by failing to purchase E than the marginal high attention type, and the social welfare gains from moving the marginal low-attention type to the energy efficient good are larger. At some levels of a policy,

b_H will be negative while b_L is positive: moving the marginal high-attention type to the efficient good will reduce welfare, while moving the marginal low-attention type will still increase welfare. The implication is that other things equal, a marginal increase in a policy n produces larger social welfare gains when D'_L is large relative to D'_H, i.e. to the extent that the inattentive types are more elastic to the policy. In Sections 5 and 6, we will examine more closely the importance of targeted policies.

4 Optimal Policy in the Vehicle Market

We have shown that in theory, inattention means that an energy tax may improve consumer welfare and that some subsidy for energy efficient durable goods is optimal. In practice, how large is this Internality Dividend from Energy Taxes, and how large should the optimal product subsidy be? In this section, we calibrate the magnitudes of our theoretical results in a simulation model of the automobile market. We first set up the simulation by detailing the supply side of the model, the choice set, and the calibration of demand parameters. We then present simulation results.

4.1 Setup

The model of the supply side is straightforward. We assume a perfectly competitive market, meaning that prices equal marginal costs. We also assume a fixed choice set, meaning that we abstract away from technological change. While markups and investments could in principle respond differently to different tax policies, they are not part of our theoretical arguments about consumer choice and optimal taxation, and endogenous changes to product offerings are particularly difficult to model credibly.

Our choice set is the set of model year 2007 new cars and trucks.[5] Models j are defined at the level of a manufacturer's model name, such as the "Honda Civic" or "Ford F-150." There are a total of 292 models in the choice set. As in the theoretical analysis, we model that there is no substitution between the new vehicle market and an outside option: a consumer will buy a new vehicle in the counterfactuals if and only if he actually did buy a new vehicle in 2007. Table 1 presents an overview of the choice set and simulation assumptions.

Vehicle prices p_j are from the JD Power and Associates "Power Information Network," a network of more than 9,500 dealers which collects detailed data on about one third of U.S. retail auto transactions. Each model's price is the mean of the final transaction price across all sales, including any customer cash rebate received from the manufacturer or dealer. If the buyer traded in a used

[5] More precisely, this is the set of 2007 new cars and trucks that have fuel economy ratings from the U.S. Environmental Protection Agency. We exclude vans as well as the following ultra-luxury and ultra-high performance exotic vehicles: the Acura NSX, Audi R8 and TT, Chrysler Prowler and TC, Cadilliac Allante and XLR Roadster, Chevrolet Corvette, Dodge Viper and Stealth, Ford GT, Plymouth Prowler, and all vehicles made by Alfa Romeo, Bentley, Ferrari, Jaguar, Lamborghini, Maserati, Maybach, Porsche, Rolls-Royce, and TVR.

vehicle, the new vehicle's price is further adjusted for the difference between the negotiated trade-in price and the trade-in vehicle's actual resale value. Market shares are from the National Vehicle Population Profile, a comprehensive national database of vehicle registrations obtained from R.L. Polk. Energy intensity e_j is the inverse of the U.S. Environmental Protection Agency (EPA) miles per gallon (MPG) fuel economy ratings. Different submodels within a model - for example, the manual vs. automatic transmission Honda Civic - may have different energy intensities, so we use each model's sales-weighted average energy intensity.

As in the theoretical model, the policymaker has two instruments, an energy tax and a product subsidy, and the government maintains a balanced budget through lump sum transfers. In this context, the "energy tax" can be thought of as a gasoline tax. Given that the choice set includes many models with many different energy intensities, the "product subsidy" now takes the form of an "energy intensity tax" τ_p that scales linearly in each model's energy intensity, increasing purchase price by amount $\tau_p e_j$. As in the theoretical model, because there is no substitution to an outside option, this energy intensity tax can equally be interpreted as an "MPG Subsidy" for energy efficient vehicles, a "feebate" that combines a fee on low-MPG vehicles with a rebate for high-MPG vehicles, or an average fuel economy standard that imposes a relative shadow cost on the sale of low-MPG vehicles.

The most uncertain parameters in the simulations are the magnitudes of the internalities, if any, and externalities. We assume that γ has a triangular distribution with mean of 0.75 and support $[0.5, 1]$. We assume that the marginal damages from uninternalized externalities ϕ from gasoline use are \$0.18 per gallon. This reflects a marginal damage from carbon dioxide emissions of \$20 per metric ton, as estimated by the U.S. Government Interagency Working Group on Social Cost of Carbon (2010). We use a pre-tax gasoline price c_g of \$3 per gallon.

We model consumers with the same utility functions as in the theoretical model, with three changes. First, we add heterogeneous preferences for different models. These preferences enter through a model-level mean utility shifter ψ_j and a consumer-by-model unobserved utility shock ϵ_{ij}. In reality, some models are more popular than others. We capture this by calibrating the mean utility shifters ψ_j such that the baseline simulated market shares equal the observed 2007 market shares. In reality, consumers' idiosyncratic preferences are often correlated within vehicle classes: some consumers have large families and prefer minivans, while rural consumers often prefer pickup trucks, and others are in the market only for sedans. To capture this, we assume that the utility shocks ϵ_{ij} have a distribution that gives nested logit substitution patterns, where the nests are nine vehicle classes defined by the U.S. EPA: pickups, sport utility vehicles, minivans, two-seaters, and five classes of cars (mini-compact, sub-compact, compact, mid-size, and large).

The second change to utility is that we add a term η which scales consumers' relative preferences for the numeraire good. The parameter η is calibrated such that the mean own-price elasticity of demand across all models is -5. This value was chosen to be consistent with the mean own-price elasticity estimated by Berry, Levinsohn, and Pakes (1995, Table V).

Third, we impose a Constant Relative Risk Aversion functional form on $u(m-\theta)$. We calibrate the parameters such that the price elasticity of demand at the mean VMT is -0.15, which is in the range of recent empirical estimates.[6] The mean θ is calibrated such that the average VMT over a potential 25-year vehicle lifetime is 236,000, which matches observed odometer readings from the National Household Travel Survey. We translate this undiscounted sum over a potential lifetime to a discounted sum over an expected lifetime by multiplying by a scaling factor $\Lambda \approx 0.436$, which accounts for observed vehicle scrappage probabilities and applies a six percent annual discount rate. See Appendix II for additional details.

After these modifications, we now have a modification of the utility function in Equation (1). The utility that consumer i experiences from purchasing product j, choosing optimal utilization m^*_{ij}, and receiving a transfer T is:

$$\eta\left\{Y_i + T - p_j - \Lambda p_g m^*_{ij} e_j\right\} + \Lambda u(m^*_{ij} - \theta_i) + \psi_j + \epsilon_{ij} \tag{12}$$

Notice that the term in brackets is consumption of the numeraire good: the amount of money from income Y_i and transfers T that the consumer has left over after purchasing the durable good and paying for gasoline. The three terms on the right represent the utility that the consumer derives from owning and using the vehicle.

As in Section 3, consumers with $\gamma_i \neq 1$ do not necessarily choose the vehicle that maximizes experienced utility. Instead, they choose vehicle j over vehicle k if and only if the perceived benefits are larger than the perceived relative costs:

$$\gamma_i\left[u(m^*_{ij}) - u(m^*_{ik})\right] + \left[\left(\psi_j + \epsilon_{ij}\right) - \left(\psi_k + \epsilon_{ik}\right)\right] > \eta\left[(p_j - p_k) + \gamma_i \Lambda p_g\left(m^*_{ij} e_j - m^*_{ik} e_k\right)\right] \tag{13}$$

To calculate welfare effects, we follow the Allcott and Wozny (2011) approach to calculating consumer surplus in logit models when consumers misoptimize. In brief, the approach exploits the fact that experienced utility can written as the difference between a decision utility function, which represents a function that the consumer acts as if he is optimizing, and the internality, which captures the magnitude by which the consumer misoptimizes. Decision consumer surplus is the integral over consumers of decision utility, which can be calculated using the nested logit version of standard discrete choice consumer surplus formulas from Small and Rosen (1981). The total internality is simply the sum over consumers of the internality. The change in experienced consumer welfare W_0 is the change in decision consumer surplus minus the change in the total internality. Interested readers can refer to Allcott and Wozny (2011) for formal details.

[6] Hughes, Knittel, and Sperling (2007) find that between 2001 and 2006, this elasticity was between -0.034 and -0.077. Small and Van Dender (2007) estimate that between 1997 and 2001, this elasticity was -0.022. Using data from California between 2001 and 2008, Gillingham (2010) estimates a short-run elasticity of -0.15 to -0.2.

4.2 Simulation Results

Table 2 presents simulation results. We simulate seven cases. Case 1 is the base equilibrium with no product tax or additional energy tax. The average new vehicle sold in 2007 has harmonic mean fuel economy 19.9 MPG. It will be driven 153,568 miles over its lifetime given observed scrappage probabilities, and as a result will emit 67.2 metric tons of CO2. The present discounted value of lifetime fuel costs for the average vehicle is $15,424.

Cases 2 and 3 assume that there are uninternalized externalities at $\phi = \$0.18$ per gallon, but that there are no internalities, i.e. that $\gamma = 1$ for all consumers. Case 2 is the first best policy: an energy tax at $\tau_g = \phi$. Case 3 applies the product subsidy that abates the same amount of carbon dioxide emissions as the first best policy in Case 2.

Cases 4-7 assume that there are both uninternalized externalities and inattention, using the triangular distribution of γ with mean 0.75. Case 4 mirrors Case 2 by applying an energy tax at $\tau_g = \phi$. Case 5 is the combination of product subsidies and energy taxes that maximize social welfare. We call this socially-optimal combination of uniform product subsidies and energy taxes the "second best." Case 6 is the product subsidy that maximizes social welfare when the energy tax is set at exactly $\tau_g = \phi$. Case 7 is the social optimum, or "first best." This could be generated by a combination of an energy tax at the level of the externality and individual-specific product subsidies that exactly correct for each individual's level of internality.

Before continuing to the core results, it is worth highlighting the importance of studying inattention. This can be seen by comparing the simulated welfare losses from externalities versus internalities. The welfare losses from externalities alone are the social welfare gains from the first best policy in Case 2: $5.48 per vehicle. The welfare losses from internalities and externalities combined are the social welfare gains from the first best policy in Case 7: $58 per vehicle. Intuitively, the additional welfare losses from inattention are so large because uninternalized carbon externalities are assumed to be $0.18 cents per gallon, or about six percent of gasoline costs, while the average inattention is assumed to be $\overline{\gamma} = 0.75$, which leaves 25 percent of gasoline costs uninternalized into product choices. Both sources of inefficiency act on the extensive margin the same way, by inducing consumers to buy vehicles that have lower fuel economy than in the social optimum, but under these parameter assumptions, inattention generates larger allocative distortions and therefore much larger welfare losses. Economists have extensively studied optimal policy under externalities. Based on these potential welfare consequences, internalities seem to merit similarly extensive study, both theoretical and empirical.

4.2.1 The Internality Dividend from Energy Taxes

Case 2 of Table 2 illustrates the traditional Pigouvian result that when externalities are the only market failure, the energy tax at the level of marginal damages reduces consumer welfare. Of course, social welfare increases from baseline by an NPV of $5.48 over the life of each new vehicle sold. However, this change in social welfare is the sum of the change in consumer welfare and the externality reduction. The externality reduction is worth \$10.84 per vehicle,[7] while consumer welfare decreases by \$5.37 per new vehicle. Aggregated over the 16 million vehicles sold in a typical year, the consumer welfare losses from Pigouvian energy taxes are \$86 million.

Case 4 shows how adding internalities to the model reverses this traditional result. The addition of the energy tax helps to reduce the pre-existing allocative distortion from inattention, increasing consumer welfare by \$4.90 per vehicle sold. Thus, the energy tax abates carbon while *increasing* consumer welfare by \$6.60 per metric ton of carbon dioxide abated. Aggregated over all new vehicles sold, a Pigouvian tax increases consumer welfare by \$78 million per year the policy is in place.

Figure 4 presents the gains in consumer and social welfare at different levels of the energy tax, under the assumption of inattentive consumers and with a product tax constrained to zero. The energy tax that maximizes consumer welfare is \$0.19 per gallon, which coincidentally is very close to the assumed level of marginal damages. Any energy tax below about \$0.38 per gallon increases consumer welfare. The social welfare-maximizing energy tax is of course larger than the consumer welfare-maximizing energy tax, as the former is set to correct distortions from externalities as well as internalities. This social-welfare maximizing level is about \$0.40 per gallon. Not coincidentally, this is slightly above the point at which a marginal increase begins to decrease consumer welfare. To see the intuition for this, consider the first order condition: the energy tax that maximizes social welfare is such that a marginal increase has zero effect on the sum of externality damages and consumer welfare.

4.2.2 The Internality Rationale for Product Taxes

A comparison of Cases 2 and 3 in Table 2 illustrates the traditional Pigouvian result that when externalities are the only market failure, the product subsidy is a highly inefficient substitute for the Pigouvian energy tax. Because marginal and average abatement costs increase in the amount of carbon dioxide abated, our comparison between the two policies must hold total abatement constant. The product subsidy that generates the same carbon dioxide abatement as the first best in Case 2 is \$67,806 per gallon per mile (GPM). To put this in perspective, a 20 MPG vehicle, such as a Subaru Outback Wagon, uses 0.05 gallons per mile, while a 25 MPG vehicle, such as a Toyota Corolla, uses 0.04 GPM. This τ_p therefore implies a relative price increase of \$678 for the 20 MPG vehicle. At this level of the product subsidy, the consumer welfare loss is \$23.30 per vehicle. This is so large that despite the gains from externality reduction, the change in social welfare is actually negative. While a smaller product tax could abate less carbon with smaller consumer welfare losses and thus generate positive social welfare gains, a smaller energy tax could still generate that smaller

[7] Intuitively, the reason why this is small relative to total lifetime gasoline costs is that the assumed carbon externality is only six percent of gasoline costs. By comparing the "Resulting Allocations" in Cases 1 and 2 of Table 2, we see that an increase in retail gasoline prices of \$0.18 per gallon does not cause a large change in either the average fuel economy of vehicles sold or the amount that they are driven.

amount of abatement much more efficiently.

Case 5 shows how adding internalities to the model reverses this traditional result that product subsidies are highly inefficient. In Case 5, we search for the combination of energy tax and product subsidy that maximizes social welfare. The optimal level of the product subsidy is \$81,261 per GPM. Using our example pair of vehicles from above, this implies a relative price increase of \$813 for the 20 MPG Subaru Outback compared to the 25 MPG Toyota Corolla.

In Proposition 3, we showed that as consumers become more inattentive, the socially-optimal energy tax remains contant and the socially-optimal product subsidy increases. Figure 5 illustrates a similar comparative static in the vehicle market simulations. This figure presents the socially-optimal combination of energy taxes and product subsidies as the average attention $\overline{\gamma}$ varies from 0.5 to 1.0. In these simulations, the support of γ is $[\overline{\gamma}/2, 1]$, which means that when $\overline{\gamma} = 1$, all consumers are fully rational. This also means that when $\overline{\gamma} = 0.75$ in the middle of the figure, the support of γ is $[0.5, 1]$ as in Table 2, and the socially-optimal tax and subsidy levels match Case 5.

Figure 5 shows that as the average attention $\overline{\gamma}$ decreases from 1 to 0.5, the optimal product tax increases from zero to approximately \$170,000 per GPM. Using our example vehicles from above, this latter level of τ_p increases the relative price of the 20 MPG Subaru Outback relative to the 25 MPG Toyota Corolla by \$1700. The key takeaway from this figure is that the optimal product tax is approximately linear in the average level of inattention, and the optimal energy tax is approximately equal to the level of the externality. This illustrates how these two market failures require two instruments. The energy tax primarily targets the externality, while the product subsidy primarily targets the internality.

Notice, however, that the optimal energy tax is not exactly equal to the externality ϕ: in Figure 6 it takes a slightly inverted U shape in $\overline{\gamma}$. When all consumers are rational, $\tau_g = \phi = \$0.18$ per gallon. In Case 5, when $\overline{\gamma} = 0.75$, the optimal gas tax is \$0.19 per gallon, and as $\overline{\gamma}$ decreases further, the optimal energy tax decreases slightly below marginal damage ϕ. Why is this the case?

To see the intuition, it is useful to contrast the vehicle market with the assumptions for Proposition 5. In that Proposition, we showed that in a world without heterogeneous product preferences ϵ_{ij} and with a homogeneous inattention parameter γ, the first best can be obtained by setting an energy tax equal to the externality and a product subsidy equal to the marginal internality. However, the vehicle market simulations include heterogeneous preferences ϵ_{ij} and heterogeneous γ, which generates variation in the utilization types θ of consumers on the margins between vehicles. Because higher-utilization consumers have higher energy costs and thus larger internalities, it is optimal to target them with larger relative price changes. The product subsidy cannot do this, as it is uniform across consumers, but the energy tax can - as long as consumers are sufficiently attentive. However, as the number of highly inattentive consumers increases, the energy tax becomes increasingly poorly targeted: it affects the product choices of the more attentive types, but not the least attentive types. Thus, as we approach the far left of the graph, the optimal tax combination eventually involves a smaller energy tax and increasing reliance on the product tax.

Depending on other parameter assumptions, when there is heterogeneity in γ but a large number of highly inattentive consumers, the optimal tax combination may actually involve a *negative* energy tax in combination with a large product subsidy. This result is again generated by differential targeting of the two policy instruments: the large product subsidy is used to reduce extensive margin distortions for low-γ consumers, while the negative energy tax is used to correct the extensive margin distortions that the large product subsidy causes for high-γ consumers. Because a well-targeted correction of the extensive margin misoptimization is relatively important from a welfare perspective and because utilization is fairly inelastic, it does not matter as much that the negative energy tax distorts utilization.

Whether the socially-optimal energy tax is above or below marginal damages depends on the joint distribution of γ, θ, and ϵ. Unfortunately, given that the average γ is difficult to infer empirically, one can expect that inferring the distribution of γ would be even more difficult. Thus, the policymaker may want to restrict attention to a "heuristic second best" policy in which $\tau_g = \phi$ and τ_p maximizes social welfare under the assumption of a homogeneous $\gamma = 0.75$. This is Case 6 in our Table 2. Importantly, the product subsidy, energy tax, and social welfare gains are nearly identical between the optimal second best policy in Case 5 and the heuristic second best policy in Case 6. Furthermore, in a series of alternative simulation runs with increasingly wide assumed distributions of γ, we found that the heuristic second best policy never does worse than achieving 98 percent of the welfare gains from the optimal second best policy. This means that for the purposes of determining the socially-optimal energy tax and product subsidy, the mean level of inattention in the population is the parameter of empirical interest, and the distribution of γ is actually of little incremental value. For other purposes, however, both the heterogeneity of γ and the elasticity of different γ-types to a product subsidy can be important, as we will soon see.

4.2.3 The Welfare Effects of Heterogeneity

In Proposition 6, we showed that as the heterogeneity in γ increases, the second best combination of product subsidies and energy taxes leaves an increasingly large remaining difference between the second best and first best level of welfare. Figure 6 illustrates this in the vehicle market simulations by plotting the welfare gains relative to baseline from the second best policies vs. first best allocations. In this figure, heterogeneity is increased by holding the peak of the triangular distribution of γ constant at 0.75 but increasing the distribution's halfwidth while truncating at support $[0, 1]$. As the heterogeneity in γ increases, the optimal combination of energy tax and product subsidy performs worse and worse relative to what is theoretically possible.

Why is this important? It means that under the reasonable supposition that consumers misoptimize in different ways - or that some consumers don't misoptimize at all - a policymaker who relies on even a perfectly-calibrated product tax could actually do much better. Ideally, the policymaker would have available other instruments that preferentially target inattentive consumers and can be used without perfect information about which consumers are making different types of mistakes.

At a minimum, it should at least be the case that the product subsidy is not for some reason more likely to affect consumers who appear to be more attentive, an issue that we explore in the next section.

5 The Targeting of Energy Efficiency Rebates

We have seen, both theoretically and in the simulated vehicle market, the importance of targeted policies. In practice, are energy efficiency subsidies well-targeted, neutral, or poorly targeted? We consider this question in the context of a North American electric utility that runs a set of highly-regarded energy efficiency programs. We examine the characteristics of households that take up energy efficiency subsidies and ask whether these are likely to be correlated with inattention to energy costs or other sources of inefficiency when investing in energy-using durables.

Notice that this exercise has two major limitations. First, while we ideally would measure how policies affect consumers with varying attentiveness to energy costs, this is not directly observed. Instead, we examine demographic characteristics that may be correlated with inattention. Second, while it is straightforward to characterize the average subsidy adopter, we will not be able to identify the marginal subsidy adopter. However, the average characteristics of subsidy adopters are still of interest, as these averages can be thought of as the integral over the consumers who are marginal at all lower levels of the subsidy.

The utility we consider, like many others, offers loans and rebates for energy efficient appliances, home improvement, and heating, ventilation, and air conditioning equipment. For example, homeowners installing new central air conditioning units that exceed a certain energy efficiency rating are eligible to receive purchase price subsidies of \$400 to \$1100 and subsidized loans of \$5000 to \$30,000. We obtained data on a sample of about 84 thousand households in the utility's service area. About 8,700 households, or one-tenth of the sample, receives a subsidy for purchasing some energy efficient good between 2007 and 2010.

We test whether subsidy recipients differ from non-recipients on several predetermined characteristics. The first is an indicator variable for whether the household received a loan or rebate to install solar panels before March 2008. The second is an indicator variable for whether the household had enrolled in the utility's "green pricing" program before March 2008. Households in this program voluntarily agree to pay a fixed additional monthly amount between \$3 and \$10 to support green energy produced by wind, hydroelectric, solar, and biomass facilities. These measures of the household's interest in green energy are plausible proxies for attentiveness to energy costs.

In addition to inattention, other economic justifications have been offered for energy efficiency subsidies. One is that consumers may be credit constrained and thus unable to pay more upfront for energy efficient goods that would save them money in the long run. A second is that landlords have a disincentive to invest in energy efficiency because renters may not be able to perfectly observe these investments, meaning that the investment might not be fully recovered via increased rents (Davis 2009, Gillingham, Harding, and Rapson 2010). We therefore examine whether subsidy recipients are less wealthy and are thus more likely to be subject to credit constraints, and whether the homes are rentals or owner-occupied.

Table 3 presents the averages of each characteristic in the populations of subsidy recipients and non-recipients, as well as a test of whether recipients differ from non-recipients. The results are clear: subsidy recipients are more likely to be "green" households that already are enrolled in green pricing and have taken up a solar panel rebate or loan. Recipients are also more likely to be wealthier and live in larger houses, which suggests that credit constraints are less likely to bind. Finally, although only a small share of the sample population are renters, subsidy recipients are even less likely to be renters than non-recipients. Each of these results suggests that these subsidies are poorly targeted toward consumers subject to the inefficiencies that justify the subsidies.

In particular, it appears that instead of encouraging inattentives to be more energy efficient, these subsidies incentivize environmentalists to be more green. What models could explain this? First, the attention parameter γ could be correlated with awareness about energy efficiency investments and subsidies. The fact that only one in ten households takes up a subsidy over the four-year period suggests that knowledge of their availability is not widespread. Takeup rates for rebates from electric utilities in many other areas of the country are much lower. It seems plausible that the exact kinds of households that do not place much weight on reducing energy costs would not seek out information on how to save energy, and might have other similarly unaware households in their social networks.

Even if all consumers were aware of the goods and their relative prices, a second plausible model is that the attention parameter γ is correlated with "rational" preferences for energy efficient goods. Hybrid vehicles and other relatively-expensive energy efficiency investments appeal strongly to "green" consumers with a taste for signaling environmentalism or privately contributing to a public good (Kahn 2007). While these goods have low market shares, both average and marginal consumers will have especially "green" preferences, and if these preferences are positively correlated with γ, then both average and marginal consumers could be relatively attentive.

Of course, these data are only suggestive that the welfare impacts of these subsidies may be compromised by poor targeting. In the next section, we formalize this idea with some explicit calculations.

6 The Lightbulb Experiment

In this section, we provide a concrete example of the welfare importance of targeting, using a randomized experiment with buyers of energy efficient lightbulbs. We first motivate the experiment, then give a conceptual overview of our goals in designing the experiment, then detail the experimental design and descriptive statistics, then present empirical results, and finally carry out the welfare calibration.

6.1 Motivation: Why Lightbulbs?

Why are lightbulbs interesting? Lighting accounts for about nine percent of household electricity use (U.S. EIA 2005), which adds up to about $11 billion per year. There are two major technologies: incandescent lightbulbs and compact fluorescent lightbulbs (CFLs). Incandescents are cheaper, but they are very inefficient at converting electricity into light: about 90 percent of the electricity that an incandescent bulb consumes is converted into heat.[8] A basic 60 watt incandescent lightbulb costs about $0.50 but consumes more than $5 in electricity over its 1000-hour lifetime. Meanwhile, a basic 60 watt equivalent CFL costs about $2 but uses only one-quarter the electricity of the incandescent. If all American consumers were to switch from incandescents to CFLs, energy costs would drop by many billions of dollars annually.

Because the energy cost savings of the CFL quickly pay back the incremental purchase price, it may be puzzling why their market shares are low: only 11 percent of residential light sockets had CFLs in 2008 (U.S. DOE 2009). One potential explanation is that consumers do not want CFLs: the light quality is different, CFLs often take some time to fully turn on, and if they break, they can release mercury. A second potential explanation is that consumers are misoptimizing when they buy incandescents: we may be unaware of or inattentive to the potential energy cost savings from CFLs, or we may systematically underestimate these savings. The idea that consumers would be better off without the option to buy incandescents is popular, and as a result, California will ban incandescents by 2018, and some kinds of incandescents will be banned nationwide beginning in 2012. A number of other countries have begun or will soon begin to ban incandescents, including Argentina, Australia, Brazil, Canada, China, Cuba, the European Union countries, Israel, Malaysia, Russia, and Switzerland. In 2008, there were 71 federal, state, and local programs in the U.S. that promoted or subsidized CFLs, at a total cost of $175 million (U.S. DOE 2010). Lightbulbs are thus an excellent case study of the issues in this paper.

6.2 Conceptual Overview

The lightbulb experiment will be used to calibrate a welfare analysis that resembles Figure 3 from earlier in the paper. If some consumers are more inattentive than others, a product subsidy will trade off gains to the less attentive types with distortions to the more attentive types. As we have seen, different attention types might have different relative elasticities to product subsidies. The goal of the experiment was to show in a practical setting how much these relative elasticities might affect welfare.

Recalling Equation (11), one needs to know three parameters to calibrate a welfare analysis with two attention types and two goods. First, one needs α, the share of consumers of the low-γ versus high-γ type. Second, one needs D'_L and D'_H, the slopes of the demand curves for each type.

[8] In fact, incandescents are so cheap and so good at producing heat that for almost 50 years, the toy manufacturer Hasbro produced an Easy-Bake Oven that used an incandescent lightbulb to heat food.

Third, one needs to know the marginal social benefits b_L and b_H from moving consumers of each attention type to the energy efficient good. This depends crucially on the inattention parameter γ.

A randomized field experiment with lightbulb buyers in a retail store affords us a sensible, albeit imperfect, opportunity to infer these parameters. Categorizing consumers as more or less attentive is especially difficult given that a purchase decision or survey response could simply reflect heterogeneous preferences for different goods. We decided to ask consumers to tell us the one or two most important factors in their purchase decision and categorize them as more attentive to energy costs if their response has to do with "energy," "energy efficiency," or "energy costs." Inferring demand curve slopes is more straightforward in an experimental setting: we are able to experimentally vary the amounts of rebate coupons given to each consumer.

Of course, inferring an inattention parameter γ is also quite challenging. Here we propose a different approach than what has been done in prior literature: we carry out an informational intervention which should leave the treated group correctly informed about and fully attentive to energy costs. This assumption is reasonable in situations where an informational intervention is plausibly powerful and correctly understood by all consumers that receive it, and if the intervention is purely informational, and does not involve social pressure or environmental messaging. Under the assumption that the consumers treated by the informational intervention have $\gamma = 1$, the average γ in the absence of the intervention can be inferred from the intervention's effect on the market share of the energy efficient good.

6.3 Experimental Design

To implement the experiment, we partnered with a large nationwide home improvement retailer that sells upwards of 50 million lightbulb packages each year. Between July and November 2011, we sent research assistants (RAs) to four stores, one in Boston, two in New York, and one in Washington, D.C. The RAs approached customers in the stores' general purpose lighting areas, which stock incandescents and CFLs that are substitutable for the same uses. Customers who consented were given a brief survey via iPad in which they were asked, among other questions, the most important factors in their lightbulb purchase decision and the amount of time each day they expected these lightbulbs to be turned on. At no time did the survey bring up energy costs. Respondents were then randomized into a two-by-two matrix of experimental conditions that included an Information treatment and a CFL Rebate treatment.

The iPad randomized half of respondents into the Information Treatment group. This group was given personally-tailored information on the energy costs. The iPad would display the annual energy costs for the bulbs in the package, given the respondent's estimated usage per day. It also displayed the total energy cost difference over the bulb lifetime and the total user cost, including energy cost and purchase price. Figure 7 is an example of the iPad information treatment screen. The RAs would interpret and discuss the costs with the customer but was instructed not to advocate for a particular type of bulb and to avoid discussing any other issues such as mercury content or environmental benefits. A typical informational intervention lasted about three minutes, and the RAs report that the information was well understood. The Information Control group did not receive the information intervention.

At the end of the survey and potential information intervention, the RAs gave respondents a coupon in appreciation for their time. The iPad randomized respondents into either the Rebate Control group, which received a coupon for 10 percent off all lightbulbs purchased, or the Rebate Treatment group, which received the same 10 percent coupon plus a second coupon valid for 30 percent off all CFLs purchased. Thus, the Rebate Treatment Group had an additional 20% discount on all CFLs. Given that the incremental price of a typical 60 Watt bulb is $1.50, this maps to a product subsidy of $0.30 per bulb. After giving customers their coupons, the RAs were instructed to leave the area so as to avoid any potential external pressure on customers' decisions. The coupons had bar codes that allowed us to observe what each respondent purchased.

The first column of Table 4 presents descriptive statistics on the population of interview respondents. Notice in particular that 25 percent of consumers reported that energy cost was an important factor in their purchase decision. The second and third columns present differences in characteristics between treatment and control groups in the Rebate and Information randomizations, respectively. In one of the 18 t-tests, a characteristic is statistically different with 95 percent confidence, and F-tests fail to reject that the groups are balanced.

Recall that our theoretical model does not include an outside option: all consumers buy either the energy efficient or energy inefficient good. To remain consistent with and to otherwise maintain simplicity, we restrict our regression sample and welfare analysis to the set of consumers that purchase a "substitutable lightbulb," by which we mean either a CFL or any incandescent or halogen that can be replaced with a CFL. The bottom part of Table 4 shows that 77 percent of interview respondents purchased any lightbulb, and 73 percent of survey respondents purchased a substitutable lightbulb. While the treatments could in theory affect whether or not to purchase a substitutable lightbulb, Table 4 shows that in practice the percentages are not significantly different between treatment and control groups. The significance levels and interpretation of our upcoming regression results do not change when we run the regressions with the slightly larger sample of people who purchased any lightbulb or with the full sample of survey respondents.

6.4 Empirical Results

The parameters needed for the welfare analysis can be inferred from the following linear probability model:

$$\begin{aligned} 1(\text{Purchase CFL})_i \quad = \quad & \beta_1 \cdot 1(\text{Information Treatment})_i + \beta_2 \cdot 1(\text{Rebate Treatment})_i \\ & +\beta_3 \cdot 1(\text{Rebate Treatment})_i \cdot 1(\text{Energy an Important Factor})_i \\ & +\beta_4 \cdot 1(\text{Energy an Important Factor})_i + \beta_0 + \varepsilon_i \end{aligned} \tag{14}$$

In this equation, i indexes individual consumers, and $1(\cdot)$ denotes the indicator function. Table 5 presents the results. Column (1) is the exact specification above, while subsequent columns include different subsets of the right-hand-side variables. The coefficients are highly robust across specifications. Column (2) shows that the rebate increased CFL purchase probability by about 10 percent.

Column (3) shows that the information intervention had no statistically significant average treatment effect on CFL purchase probability. In fact, the standard errors are tight enough to bound the effect to being less than about 2/3 the effect of the CFL rebate, which at about $0.30 per bulb was not very large. In the welfare analysis, we show the implications of two competing interpretations of this result. First, one could interpret it to mean that all consumers have $\gamma = 1$. This is consistent with the fact that our partner retailer already has a substantial amount of easy-to-understand informational and promotional materials about CFLs in the general purpose lighting section. Second, one could interpret it to mean that consumers have an average $\overline{\gamma} = 1$, and that some consumers in the absence of the information intervention have $\gamma < 1$, while others have $\gamma > 1$. This is consistent with the idea that some consumers underestimate energy cost savings from CFLs, while others overestimate. Because the consumers that shop in home improvement stores in large east coast cities may be different than consumers that buy lightbulbs elsewhere, and because our partner retailer has better existing informational materials than supermarkets and hardware stores that also sell a large number of lightbulbs, we emphasize that this zero effect does not provide any generalizable evidence that could be used to argue for or against any nationwide regulation.

Column (4) shows that those who report that energy is an important factor in their purchase decision are just under 40 percentage points more likely to buy CFLs than than those who do not. The CFL market share for these consumers that we have categorized as "more attentive" is almost twice the CFL market share for the consumers categorized as "less attentive." While this result is consistent with our assumption that this variable can be used to categorize consumers into two attention types, it certainly does not prove our assumption, as consumers who do not report that energy is an important factor could simply have rationally stronger preferences for incandescents. For example, they could have lower marginal utility of money or feel less warm glow utility from saving energy.

Column (5) tests whether more vs. less attentive types have different elasticities to the rebate. We interact the Rebate Treatment indicator with the indicator for whether energy is an important factor in the purchase decision. Taking the point estimates literally, the "less attentive" types have a

12.9 percentage point response to the rebate, while the "more attentive" types have a 12.9-9.4 = 3.5 percentage point response. This difference has a p-value of 0.16 and is therefore not statistically significant. However, these results can be used to generate bounding cases: the minimum and maximum possible difference in demand slopes that can be admitted by the 95 percent confidence interval. These provide best and worst case scenarios for the targeting of the subsidy.

6.5 Welfare Calibration

We now use these empirical data to calibrate the optimal product subsidy and welfare effects. We focus on the subsidy that maximizes consumer welfare, which is also the subsidy that maximizes social welfare when energy is priced at social cost. It is debatable whether retail electricity prices are above or below long run marginal social cost due to various retail pricing inefficiencies, and we choose to abstract away from these issues.

Table 6 presents the welfare calculation. The parameter $\alpha = 0.25$ reflects the fact that one-quarter of survey respondents list energy as an important factor in their purchase decision. We set $\overline{\gamma} = 1$ to reflect the zero Average Treatment Effect of the information intervention. The parameters γ_L and γ_H are pinned down by two equations: that the weighted average γ is one, and that the difference in CFL purchase probabilities between the low-attention and high-attention types is about 40 percentage points, as reported in Table 5.[9] D' is the average slope of the demand curve in purchase probability per dollar; this is determined by dividing the treatment effect of the rebate (9.5 percentage points) by the average amount of the rebate per bulb ($0.30). The optimal subsidy is determined by an analogue to Equation (9).

Columns 1 and 2 of Table 6 simply remind us that the optimal product subsidy is zero when the average marginal internality is zero. When this is the case, any product subsidy will reduce consumer welfare. Column 3 shows the case where the average population internality and the slope of aggregate demand are held constant, but the difference in slopes between the less attentive versus more attentive types is as large as can be admitted by the 95 percent confidence interval of Column (5) in Table 5. (In this case, the 95 percent confidence interval actually allows upward sloping demand for the high-γ type, so we bound D'_H at zero.) This provides an upper bound on the average marginal internality, and thus an upper bound on the optimal product subsidy. This upper bound is $0.40 per bulb, just larger than the experimental rebate, and about 27 percent of the incremental price of a basic 60 watt equivalent CFL compared to a 60 watt incandescent.

Column 4 presents the opposite case to Column 3: when the difference in slopes between the more attentive versus less attentive types is as large as can be admitted by the 95 percent confidence

[9] Specifically, these equations are:

$$\alpha\gamma_H + (1-\alpha)\gamma_L = 1$$
$$p_g m(e_I - e_E) \cdot (\gamma_H - \gamma_L) \cdot D' = \Pr(PurchaseCFL|\gamma = \gamma_H) - \Pr(PurchaseCFL|\gamma = \gamma_L) \approx 0.4.$$

We calibrate p_g at $0.10 per kilowatt-hour, which is the national average retail electricity price. We set m at 1000 hours, the typical life of an incandescent bulb. The parameters e_I and e_E are 60 and 15 Watts, respectively.

interval. This would imply that the high-γ types are more responsive to the rebate than the low-γ types, and thus that the average marginal consumer misoptimizes by buying *too many* CFLs. In this case, the optimal CFL subsidy is *negative*: the policymaker would want to subsidize the energy inefficient good. While this causes welfare losses for low-γ consumers, these are outweighed by welfare gains to the high-γ types.

Certainly, this analysis is only suggestive: the parameters apply only for our limited experimental population, and we hope that future work will develop better approaches to determining which consumers, if any, are misoptimizing, and by how much. This should not obscure the fact that this calibration provides guidance on and examples of the parameters necessary to set optimal policy and perform welfare analysis.

This specific calibration also has several potentially-provocative takeaways. First, Column 1 of Table 6 makes the basic point that while an energy efficiency subsidy causes a decrease in energy use, there is no evidence that it increases welfare unless there is evidence of some inefficiency. Although this is unsurprising for economists, it is worth highlighting because it flows directly from one interpretation of the zero effect of our informational intervention. Second, Column 3 illustrates how a subsidy could have positive welfare effects even if the average consumer is not inattentive. What matters is the average marginal internality, not the average internality.

Third, Column 4 considers a plausible case in which overattentive types are more elastic to the subsidy. As suggested by the results of the previous section, the kinds of people that overestimate energy costs may be exactly the kinds of people who are more responsive to changes in the relative price of energy efficient goods. If this is true, it may be optimal to tax energy efficient goods instead of subsidizing them. Cost benefit analyses of energy policies typically focus on how much energy is saved. The welfare calibration of the lightbulb experiment illustrates the importance of understanding *who* is saving the energy.

7 Conclusion: Mechanisms for Behavioral Targeting

Many analysts and policymakers believe that consumers misoptimize in ways that cause us to underinvest in energy efficient durable goods. In this paper, we study optimal policy design when some consumers are inattentive to energy costs. We show that inattention reverses two traditional results from a world where externalities are the only market failure: there is now an Internality Dividend from Energy Taxes and an Internality Rationale for Product Subsidies. However, heterogeneity across consumers in the magnitudes of their internalities means that policies that preferentially target misoptimizers have larger welfare gains. We present a case study of this fact using a field experiment with lightbulb buyers, as well as some suggestive evidence that existing energy efficiency subsidies may be poorly targeted. A basic takeaway is that the average marginal internality is a crucial object for optimal policy design and welfare analysis.

How can policies be designed to target inattentive consumers? We present four economically-motivated ideas and examples of existing policies that implement them. First, in the spirit of Akerlof's (1978) discussion of targeted social programs, policymakers can use *behavioral tagging*: limiting eligibility to individuals with observable characteristics correlated with misoptimization. For example, many utilities mail energy conservation reports to some of their residential customers, but only send them to homeowners with relatively high energy use, who are more likely to be inattentive or poorly informed (Allcott 2011b). Utilities could also limit subsidies to first-time participants in energy efficiency programs if they believe that repeat participants are more likely to be fully informed and attentive.

Second, policymakers can use *behavioral screening*: offering incentives that misoptimizing consumers are more likely to adopt. For example, some energy efficiency programs subsidize the cost of weatherization investments equally for all households, while others make the household's subsidy a function of estimated energy savings. The latter structure is better targeted at inattentive types, as the marginal inattentive types will tend to have larger potential energy savings than the marginal attentive types.

Third, policymakers can exploit *nudges*: factors that affect misoptimizing consumers without affecting the behavior of rational consumers (Thaler and Sunstein 2008). Information provision programs such as appliance and vehicle energy use tags are one example, as these both draw attention to energy costs and inform the informed. Another example is "on-bill financing" programs, in which the utility pays part of the upfront cost of a home energy efficiency investment and amortizes that cost over several years on the homeowner's energy bills. While these have traditionally been justified as a way to alleviate credit constraints, another useful feature of on-bill financing is that it puts upfront investment costs and future energy costs into the same payment stream, eliminating the possibility that the consumer could attend differently to the two types of costs.

Fourth, electric utilities and retailers of energy-using durables often have more powerful capacity to inform or nudge consumers than the government: they can provide their own energy cost informational materials to complement any mandated information disclosure, or alternatively hide the required materials at the back of the retail floor. Firms can also direct their retail sales staff either to make extra effort to inform consumers about the energy costs of different models, or to instead focus on other attributes. A policymaker can induce firms to nudge consumers by *externalizing the internality*: implementing a tax, subsidy, or other policy instrument that inserts a correlate of the internality into firms' profit functions. One potential example of a "nudge-inducing policy" is the Energy Efficiency Resource Standard, which requires electric utilities to induce their customers to conserve a required amount of energy per year. These policies are imperfect, however, as they incentivize energy conservation from any consumer, not just from consumers that misoptimize. As we have seen, it matters not just *how much* energy is conserved, but *who* is conserving.

References

[1] Abaluck, Jason, and Jonathan Gruber (2011). "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, Vol. 101, No. 4 (June), pages 1180–1210.

[2] Akerlof, George (1978). "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review*, Vol. 68, No. 1 (March), pages 8-19.

[3] Allcott, Hunt (2011a). "Consumers' Perceptions and Misperceptions of Energy Costs." *American Economic Review*, Vol. 101, No. 3 (May), pages 98-104.

[4] Allcott, Hunt (2011b). "Social Norms and Energy Conservation." *Journal of Public Economics*, Vol. 95, No. 9-10 (October), pages 1082-1095.

[5] Allcott, Hunt (2012). "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." Working Paper, New York University (January).

[6] Allcott, Hunt, and Michael Greenstone (2012). "Is there an Energy Efficiency Gap?" *Journal of Economic Perspectives*, Vol. 26, No. 1 (Winter), pages 3-28.

[7] Allcott, Hunt, and Sendhil Mullainathan (2010). "Behavior and Energy Policy." *Science*, Vol. 327, No. 5970 (March 5th).

[8] Allcott, Hunt, and Nathan Wozny (2011). "Gasoline Prices, Fuel Economy, and the Energy Paradox." Working Paper, Massachusetts Institute of Technology (May).

[9] Attari, Shahzeen, Michael DeKay, Cliff Davidson, and Wandi Bruine de Bruin (2010). "Public Perceptions of Energy Consumption and Savings." *Proceedings of the National Academy of Sciences*, Vol. 107, No. 37 (September 14), pages 16054-16059.

[10] Atkinson, Scott, and Robert Halvorsen (1984). "A New Hedonic Technique for Estimating Attribute Demand: An Application to the Demand for Automobile Fuel Efficiency." *Review of Economics and Statistics*, Vol. 66, No. 3 (August), pages 417-426.

[11] Austin, David (2008). "Effects of Gasoline Prices on Driving Behavior and Vehicle Markets." U.S. Congressional Budget Office Working Paper (January).

[12] Austin, David, and Terry Dinan. 2005. "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes." *Journal of Environmental Economics and Management*, Vol. 50, No. 3, pages 562–582.

[13] Barber, Brad, Terrance Odean, and Lu Zheng (2005). "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows." *Journal of Business*, Vol. 78, No. 6 (November), pages 2095-2120.

[14] Bento, Antonio, Lawrence Goulder, Mark Jacobsen, and Roger von Haefen (2009). "Distributional and Efficiency Impacts of Increased US Gasoline Taxes." American Economic Review, Vol. 99, No. 3 (June), pages 667-699.

[15] Bernheim, B. Douglas, and Antonio Rangel (2004). "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, Vol. 90, No. 5 (Decmeber), pages 1558-1590.

[16] Bernheim, B. Douglas, and Antonio Rangel (2005). "From Neuroscience to Public Policy: A New Economic View of Addiction." *Swedish Economic Policy Review*, Vol. 12, pages 11-56.

[17] Bernheim, B. Douglas, and Antonio Rangel (2009). "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics*, Vol. 124, No. 1 (February), pages 51-104.

[18] Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63, No. 4 (July), pages 841-890.

[19] Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman (2010). "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *Quarterly Journal of Economics*, Vol. 125, No. 1, pages 263-305.

[20] Bovenberg, Lans, and Lawrence Goulder (1996). "Optimal Environmental Taxation in the Presence of Other Taxes: General Equilibrium Analyses." *American Economic Review*, Vol. 86, No. 4, pages 985-1000.

[21] Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer (2011). "Pain at the Pump: How Gasoline Prices Affect Automobile Purchasing." Working Paper, Massachusetts Institute of Technology (May).

[22] Busse, Meghan, Jorge Silva-Risso, Florian Zettelmeyer (2006). "$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *American Economic Review*, Vol. 96, No. 4 (Septemper), pages 1253-1270.

[23] Cabral, Marika, and Caroline Hoxby (2010). "The Hated Property Tax: Salience, Tax Rates, and Tax Revolts." Working Paper, Stanford University.

[24] Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin (2003). "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism." *University of Pennsylvania Law Review*, Vol. 151, pages 1211-1254.

[25] Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence." *American Economic Review*, Vol. 99, No. 4 (September), pages 1145-1177.

[26] Dahlby, Bev (2008). The Marginal Cost of Public Funds: Theory and Applications. Cambridge, MA: The MIT Press.

[27] Davis, Lucas (2008). "Durable Goods and Residential Demand for Energy and Water: Evidence from a Field Trial." *RAND Journal of Economics*, Vol. 39, No. 2 (Summer), pages 530-546.

[28] Davis, Lucas (2010). "Evaluating the Slow Adoption of Energy Efficient Investments: Are Renters Less Likely to Have Energy Efficient Appliances?" In Fullerton, Don, and Catherine Wolfram, Eds., The Design and Implementation of U.S. Climate Policy.

[29] Davis, Lucas, and Lutz Kilian (2009). "Estimating the Effect of a Gasoline Tax on Carbon Emissions." National Bureau of Economic Research Working Paper 14685 (January).

[30] Davis, William, Levine, Mark and Train, Kenneth, with K.G. Duleep (1993). "Feebates: Estimated Impacts on Vehicle Fuel Economy, Fuel Consumption, CO2 Emissions, and Consumer Surplus." Draft Report, Lawrence Berkeley Laboratory, Berkeley, California.

[31] Diamond, Peter (1973). "Consumption Externalities and Imperfect Corrective Pricing." *Bell Journal of Economics and Management Science*, Vol. 4, No. 2 (Autumn), pages 526-538.

[32] Dreyfus, Mark, and Kip Viscusi (1995). "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." *Journal of Law and Economics*, Vol. 38, No. 1, pages 79-98.

[33] Dubin, Jeffrey, and Daniel McFadden (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica*, Vol. 52, No. 2 (March), pages 345-362.

[34] Espey, Molly, and Santosh Nair (2005). "Automobile Fuel Economy: What Is It Worth?" *Contemporary Economic Policy*, Vol. 23, No. 3 (July), pages 317-323.

[35] Finkelstein, Amy (2009). "EZ-Tax: Tax Salience and Tax Rates." *Quarterly Journal of Economics*, Vol. 124, No. 3 (August), pages 969-1010.

[36] Fischer, Carolyn, Winston Harrington, and Ian Parry (2007). "Do Market Failures Justify Tightening Corporate Average Fuel Economy (CAFE) Standards?" *The Energy Journal*, Vol. 28, No. 4, pages 1-30.

[37] Fullerton, Don, and Sarah West (2010). "Tax and Subsidy Combinations for the Control of Car Pollution." *B.E. Journal of Economic Analysis and Policy: Advances*. Vol. 10, No. 1 (January), pages 1-31.

[38] Gabaix, Xavier, and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, Vol. 121, No. 2, pages 505-540.

[39] Gallagher, Kelly Sims, and Erich Muehlegger (2008). "Giving Green to Get Green? Incentives and Consumer Adoption of Hybrid Vehicle Technology." Working Paper, Harvard University (January).

[40] Gately, Dermot (1980). "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables: Comment." *Bell Journal of Economics*, Vol. 11, No. 1 (Spring), pages 373-374

[41] Gillingham, Kenneth (2010). "Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California." Working Paper, Stanford University (September).

[42] Goldberg, Pinelopi (1998). "The Effects of the Corporate Average Fuel Efficiency Standards in the US." Journal of Industrial Economics, Vol. 46, No. 1 (March), pages 1-33.

[43] Greene, David (1983). "A note on implicit consumer discounting of automobile fuel economy: Reviewing the available evidence." *Transportation Research Part B: Methodological*, Vol. 17, No. 6 (December), pages 491-499.

[44] Greene, David (1998). "Why CAFE Worked." *Energy Policy*, Vol. 26, No. 8, pages 595-613.

[45] Greene, David (2010). "How Consumers Value Fuel Economy: A Literature Review." US Environmental Protection Agency Technical Report EPA-420-R-10-008 (March).

[46] Greene, David, Philip Patterson, Margaret Singh, and Jia Li (2005). "Feebates, Rebates, and Gas-Guzzler Taxes: A Study of Incentives for Increased Fuel Economy." *Energy Policy*, Vol. 33, No. 6 (April), pages 757-775.

[47] Gillingham, Kenneth (2010). "Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California." Working Paper, Stanford University (September).

[48] Gillingham, Kenneth, Matthew Harding, and David Rapson (2010). "Split Incentives in Residential Energy Consumption." Working Paper, Stanford University.

[49] Gillingham, Kenneth, Richard Newell, and Karen Palmer (2006). "Energy Efficiency Policies: A Retrospective Examination." *Annual Review of Environment and Resources*, Vol. 31, pages 161-192.

[50] Gillingham, Kenneth, Richard Newell, and Karen Palmer (2009). "Energy Efficiency Economics and Policy." Resources for the Future Discussion Paper 09-13 (April).

[51] Gruber, Jonathan (2002). "Smokers' Internalities." *Regulation*, Vol. 25, No. 4, pages 25-57.

[52] Gruber, Jonathan, and Botund Koszegi (2004). "Tax Incidence When Individuals are Time Inconsistent: The Case of Cigarette Excise Taxes." *Journal of Public Economics*, Vol. 88, No. 9-10 (August), pages 1959-1988.

[53] Gruber, Jonathan, and Sendhil Mullainathan (2002). "Do Cigarette Taxes Make Smokers Happier?" NBER Working Paper No. 8872 (April).

[54] Hausman, Jerry (1979). "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *Bell Journal of Economics*, Vol. 10, No. 1, pages 33-54.

[55] Hausman, Jerry, and Paul Joskow (1982). "Evaluating the Costs and Benefits of Appliance Efficiency Standards." *American Economic Review*, Vol. 72, No. 2 (May), pages 220-225.

[56] Heutel, Garth (2011). "Optimal Policy Instruments for Externality-Producing Durable Goods under Time Inconsistency." Working Paper, UNC Greensboro (May).

[57] Herrnstein, R. J., George Loewenstein, Drazel Prelec, and William Vaughan, Jr. (1993). "Utility Maximization and Melioration: Internalities in Individual Choice." *Journal of Behavioral Decision Making*, Vol. 6, pages (149-185).

[58] Hossain, Tanjim, and John Morgan (2006). "...Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis and Policy*, Vol. 6.

[59] Hughes, Jonathan, Christopher Knittel, and Daniel Sperling (2008). "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *The Energy Journal*, Vol. 29, No. 1, pages 113-134.

[60] Jacobsen, Mark (2010a). "Evaluating U.S. Fuel Economy Standards in a Model with Producer and Household Heterogeneity." Working Paper, University of California at San Diego (September).

[61] Jacobsen, Mark (2010b). "Fuel Economy, Car Class Mix, and Safety." Working Paper, University of California at San Diego (November).

[62] Jaffe, Adam, and Robert Stavins (1994). "The Energy Paradox and the Diffusion of Conservation Technology." *Resource and Energy Economics*, Vol. 16, pages 91-122.

[63] Kilian, Lutz and Eric Sims (2006). "The Effects of Real Gasoline Prices on Automobile Demand: A Structural Analysis Using Micro Data," Working Paper, University of Michigan (April).

[64] Kleit, Andrew N. 2004. Impacts of Long-Range Increases in the Corporate Average Fuel Economy (CAFE) Standard. *Economic Inquiry*, Vol. 42, No. 2, pages 279-294.

[65] Koszegi, Botond, and Paul Heidhues (2010). "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, Vol. 100, No. 5, pages 2279-2303.

[66] Krupnick, Alan, Ian Parry, Margaret Walls, Tony Knowles, and Kristin Hayes (2010). “Toward a New National Energy Policy: Assessing the Options.” Resources for the Future Report (November). http://www.energypolicyoptions.org/wp-content/uploads/reports/RFF-Rpt-NEPI_Tech_Manual_Final.pdf

[67] Langer, Ashley, and Nathan Miller (2009). "Automakers’ Short-Run Responses to Changing Gasoline Prices and the Implications for Energy Policy." Working Paper, University of California at Berkeley (September).

[68] Larrick, Richard, and Jack Soll (2008). “The MPG Illusion.” *Science*, Vol. 320, No. 5883, pages 1593-1594.

[69] Murtishaw, Scott, and Jayant Sathaye (2006). "Quantifying the Effect of the Principal-Agent Problem on US Residential Energy Use." Working Paper, Lawrence Berkeley National Laboratory (August).

[70] National Highway Traffic Safety Administration (NHTSA) (2010). "Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012-MY 2016 Passenger Cars and Light Trucks." Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis (March).

[71] Neubauer, Max, Andrew deLaski, Marianne DiMascio, and Steven Nadel (2009). "Ka-BOOM! The Power of Appliance Standards." ACEEE Report Number AO91.

[72] O’Donoghue, Ted, and Matthew Rabin (2006). "Optimal Sin Taxes." *Journal of Public Economics*, Vol. 90, pages 1825-1849.

[73] Ohta, Makoto, and Zvi Griliches (1986). "Automobile Prices and Quality: Did the Gasoline Price Increases Change Consumer Tastes in the U.S.?" *Journal of Business and Economic Statistics*, Vol. 4, No. 2 (April), pages 187-198.

[74] Parry, Ian (1995). "Pollution Taxes and Revenue Recycling." *Journal of Environmental Economics and Management*, Vol. 29, pages S64-S77.

[75] Parry, Ian, David Evans, and Wallace Oates (2010). "Are Energy Efficiency Standards Justified? Resources for the Future Discussion Paper 10-59 (November).

[76] Parry, Ian, Winston Harrington, and Margaret Walls (2007). "Automobile Externalities and Policies." *Journal of Economic Literature*, Vol. 45, pages 374-400.

[77] Pearce, David (1991). "The Role of Carbon Taxes in Adjusting to Global Warming." *The Economic Journal*, Vol. 101, No. 407 (July), pages 938-948.

[78] Pigou, Arthur (1932). The Economics of Welfare. London: MacMillan and Co.

[79] Reiss, Peter, and Matthew White (2008). "What Changes Energy Consumption? Prices and Public Pressure." *RAND Journal of Economics*, Vol. 39, No. 3 (Autumn), pages 636-663.

[80] Sallee, James (2011a). "The Taxation of Fuel Economy." Chapter 1 in Jeffrey Brown, Ed., Tax Policy and the Economy, Volume 25. Cambridge, MA: National Bureau of Economic Research.

[81] Sallee, James (2011). "Rational Inattention and Energy Efficiency." Working Paper, University of Chicago (June).

[82] Sallee, James, Sarah West, and Wei Fan (2011). "The Effect of Gasoline Prices on the Demand for Fuel Economy in Used Vehicles: Empirical Evidence and Policy Implications." Work in Progress, University of Chicago.

[83] Sawhill, James (2008). "Are Capital and Operating Costs Weighted Equally in Durable Goods Purchases? Evidence from the US Automobile Market." Working Paper, University of California at Berkeley (April).

[84] Schwartzstein, Joshua (2010). "Selective Attention and Learning." Working Paper, Dartmouth (December).

[85] Small, Kenneth, and Kurt Van Dender (2007). "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect." *The Energy Journal*, Vol. 28, No. 1, pages 25-51.

[86] Sunstein and Thaler (2003). "Libertarian Paternalism is not an Oxymoron." *University of Chicago Law Review*, Vol. 70, No. 4 (Fall), pages 1159-1202.

[87] Thaler, Richard, and Cass Sunstein (2009). Nudge. New York, NY: Penguin Books.

[88] Train, Kenneth, (1985). "Discount Rates in Consumers' Energy-related Decisions: A Review of the Literature." *Energy*, Vol. 10, No. 12, pages 1243-1253.

[89] Tullock, G (1967). "Excess Benefit." *Water Resources Research*, Vol. 3, pages 643-644.

[90] U.S. Department of Energy (DOE) (2009). "CFL Market Profile." http://www.energystar.gov/ia/products/downloads/CFL_Market_Profile.pdf

[91] U.S. Energy Information Administration (EIA) (2010). "2009 Electric Power Annual." http://www.eia.gov/cneaf/electricity/epa/epa_sum.html

[92] U.S. Environmental Protection Agency (EPA) (2010a). "Gas Guzzler Tax." http://www.epa.gov/fueleconomy/guzzler/index.htm.

[93] U.S. Environmental Protection Agency (EPA) (2010b). "New Energy Tax Credits for Hybrids." http://www.fueleconomy.gov/feg/tax_hybrid.shtml.

[94] U.S. Government, Interagency Working Group on Social Cost of Carbon (2010). "Technical Support Document: - Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866 -." http://www.epa.gov/oms/climate/regulations/scc-tsd.pdf

[95] Verboven, Frank (1999). "The Markets for Gasoline and Diesel Cars in Europe." Centre for Economic Policy Research Discussion Paper 2069 (February).

[96] Verboven, Frank (2002). "Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars." *Rand Journal of Economics*, Vol. 33, No. 2 (Summer), pages 275-323.

Tables

Table 1: Vehicle Market Simulation Overview

	Mean	*Standard Deviation*	*Min*	*Max*
Choice Set				
Number of Models	292			
Price p_j (\$)	36,615	25,086	12,038	174,541
Gallons per Mile e_j	0.054	0.011	0.022	0.084
2007 Quantity Sold	47,891	75,155	93	627,809
Energy				
Pre-Tax Gasoline Price p_g (\$ per gallon)	3			
Marginal Damage ϕ (\$ per gallon)	0.18			
Consumers				
Attention Weight γ	0.75	0.10	0.5	1
Nested Logit Substitution Parameter σ	0.6			
Mean Own-Price Elasticity	-5			
VMT Elasticity	0.15			
Baseline Lifetime Potential VMT m^*	237,220	72,870	108,400	382,840
Annual Discount Rate	6%			
Lifetime Scaling Factor Λ	0.436			

Notes: This table presents the basic assumptions for the vehicle market simulations. All dollars are real 2005 dollars.

Table 2: Vehicle Market Simulation Results

Source of Inefficiency							
Externalities Only		Yes	Yes				
Externalities and Inattention				Yes	Yes	Yes	Yes
Case	1	2	3	4	5	6	7
	Base Case	First Best: $\tau_g = \phi$, $\tau_p = 0$	$\tau_g = 0$, τ_p to Abate Same CO2 as Case 2	$\tau_g = \phi$ $\tau_p = 0$	Second Best: τ_g and τ_p to Max Social Welfare	"Heuristic Second Best": $\tau_g = \phi$, τ_p to Max Social Welfare	First Best
Policies							
Gas Tax τ_g ($/gallon)	0.00	0.18	0.00	0.18	0.19	0.18	0.18
Product Subsidy τ_p ($/GPM)	0	0	67,806	0	81,261	81,805	
Resulting Allocations							
Average MPG	19.9	19.9	20.1	19.9	20.3	20.3	20.3
Average Lifetime VMT	153,568	152,307	153,910	152,287	152,630	152,695	152,701
Average PDV of Gas Cost	15,424	16,158	15,243	16,192	15,998	15,961	15,938
Average CO2 Tons Emitted	67.2	66.4	66.4	66.5	65.6	65.6	65.5
Welfare Compared to Base							
ΔConsumer Welfare/Vehicle		-5.37	-23.3	4.9	25.4	25.7	33.4
ΔCO2 Damages/Vehicle		-10.84	-10.8	-10.1	-23.6	-23.2	-24.5
ΔSocial Welfare/Vehicle		5.48	-12.4	15.0	48.9	48.9	58.0
ΔConsumer Welfare/ton CO2		-6.82	-29.5	6.6	14.8	15.3	18.8

Notes: This table presents the results of the vehicle market simulations. All dollars are real 2005 dollars. Carbon emissions and damages are denominated in metric tons of carbon dioxide. Average Lifetime VMT accounts for vehicle scrappage probabilities over each year of the 25-year maximum life. Welfare numbers are dollars per new vehicle buyer over the life of the vehicle, discounted at 6 percent per year.

Table 3: Characteristics of Energy Efficiency Subsidy Recipients

	Subsidy Recipients	*Non-Recipients*	*Difference*
1(Green Pricing)	0.29	0.23	0.057
	(0.45)	(0.42)	(.0051***)
1(Solar Loan/Rebate)	0.0018	0.0004	0.0014
	(0.042)	(0.021)	(.0005***)
House Value ($000s)	218	214	4.3
	(141)	(147)	(1.6***)
Income ($millions)	91	85	5.4
	(42)	(42)	(0.48***)
1(Rent)	0.0035	0.01	-0.0064
	(0.0055)	(0.095)	(.0007***)
N	75,146	8,686	83,832

Notes: Columns 1 and 2 present mean characteristics for energy efficiency subsidy recipients and non-recipients, with standard deviations in parenthesis. Column 3 presents the difference in means, with robust standard errors in parenthesis. *, **, ***: Statistically significant with 90%, 95%, and 99% confidence, respectively.

Table 4: Lightbulb Experiment Descriptive Statistics

Individual Characteristics	*Experimental Population Mean*	*CFL Rebate T - C Difference*	*Information T - C Difference*
Energy an Important Factor	0.25	-0.024	0.009
	(0.43)	(0.026)	(0.026)
Expected Usage (Minutes/Day)	333.0	2.7	12.8
	(280.0)	(17.0)	(17.0)
Age	43.8	-0.3	0.7
	(11.4)	(0.7)	(0.7)
Male	0.66	0.003	0.009
	(0.47)	(0.029)	(0.029)
African American	0.16	-0.008	-0.001
	(0.37)	(0.022)	(0.022)
Asian	0.06	0.005	-0.030
	(0.24)	(0.015)	(0.014**)
Caucasian	0.66	-0.005	0.037
	(0.47)	(0.029)	(0.029)
Hispanic	0.07	0.011	0.001
	(0.25)	(0.015)	(0.015)
Middle Eastern	0.01	0.007	0.002
	(0.12)	(0.007)	(0.013)
F-Test p-Value		0.896	0.742
Number of Observations	1087	1087	1087
Regression Sample			
Purchased Any Lightbulb	0.77	0.027	0.011
	(0.42)	(0.025)	(0.025)
Purchased Substitutable Lightbulb	0.73	0.011	-0.008
	(0.44)	(0.027)	(0.027)

Notes: The first column presents means of individual characteristics in the survey population, with standard deviations in parenthesis. The second and third columns present differences in means between treatment and control groups, with robust standard errors in parenthesis. *, **, ***: Statistically significant with 90%, 95%, and 99% confidence, respectively. The bottom panel of the table shows the determination of the regression sample, which was the subset of the survey population that purchased a substitutable lightbulb.

Table 5: Lightbulb Experiment Regression Results

	I	II	III	IV	V
	(1)	(2)	(3)	(4)	(5)
Information Treatment	-.003		-.010		
	(0.033)		(0.035)		
Rebate Treatment	0.128	0.095			0.129
	(0.042)***	(0.035)***			(0.042)***
(Rebate Treatment)x(Energy an Important Factor)	-.094				-.094
	(0.067)				(0.067)
Energy an Important Factor	0.426			0.38	0.426
	(0.047)***			(0.034)***	(0.047)***
Const.	0.341	0.477	0.528	0.402	0.339
	(0.034)***	(0.025)***	(0.025)***	(0.021)***	(0.029)***
Obs.	794	794	794	794	794
R^2	0.137	0.009	0.0001	0.125	0.137
F statistic	35.494	7.173	0.082	128.21	47.384

Notes: This table presents the results of estimating Equation (14). Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90%, 95%, and 99% confidence, respectively.

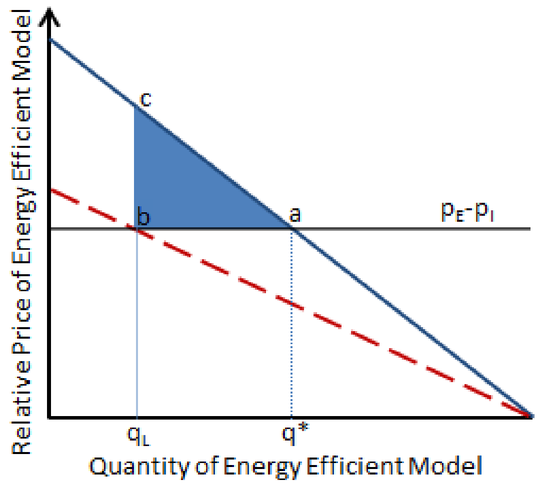
Table 6: Lightbulb Experiment Welfare Calculations

Scenario	1	2	3	4
α		0.25	0.25	0.25
$\overline{\gamma}$	1	1	1	1
γ_L		0.91	0.91	0.91
γ_H		1.27	1.27	1.27
D'_L		0.32	0.42	0.29
D'_H		0.32	0.00	0.41
D'	0.32	0.32	0.32	0.32
Optimal Product Subsidy				
Optimal CFL Subsidy ($/bulb)	0	0	0.40	-0.01
Optimal CFL Subsidy (% of Relative Price)	0	0	0.27	-0.01
Subsidy Effect on CFL Purchase Probability				
Low-γ Consumers		0	0.169	-0.003
High-γ Consumers		0	0.000	-0.004
Average Consumer	0	0	0.127	-0.003
Subsidy Effect on Consumer Welfare				
Low-γ ($/Consumer)		0	0.101	-0.001
High-γ ($/Consumer)		0	0.000	0.004
Average ($ /Consumer)	0	0	0.076	0.0003
ΔConsumer Welfare/million bulbs	0	0	76,000	315

Notes: This table presents optimal CFL subsidies and welfare effects under different assumptions for how different attention types respond to the subsidy.

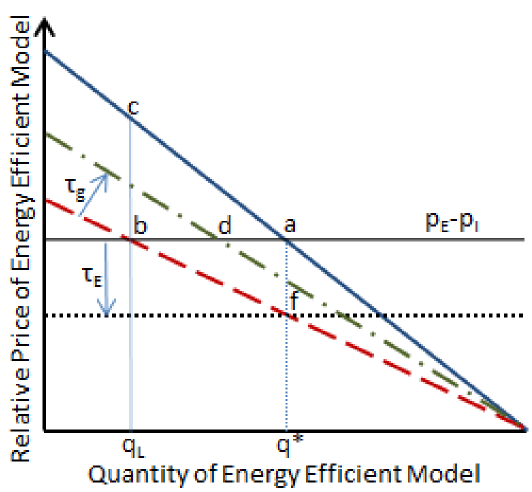
Figures

Figure 1: Baseline Equilibrium Under Inattention



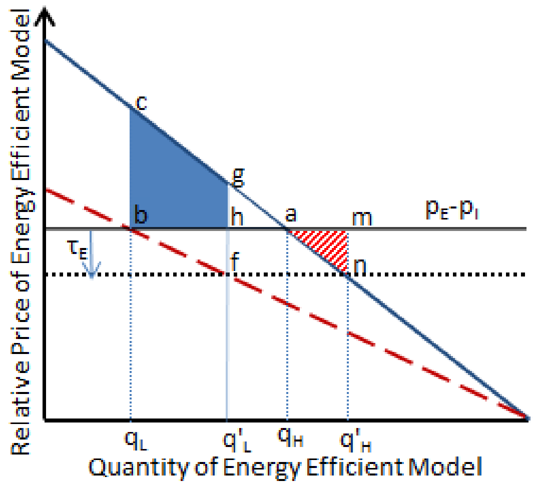
Notes: The solid blue line on this figure is the demand curve for the energy efficient good under the rational model. The dashed red line is the demand curve under inattention. Triangle abc is the consumer welfare loss from inattention.

Figure 2: Equilibrium with Energy Taxes or Product Subsidies



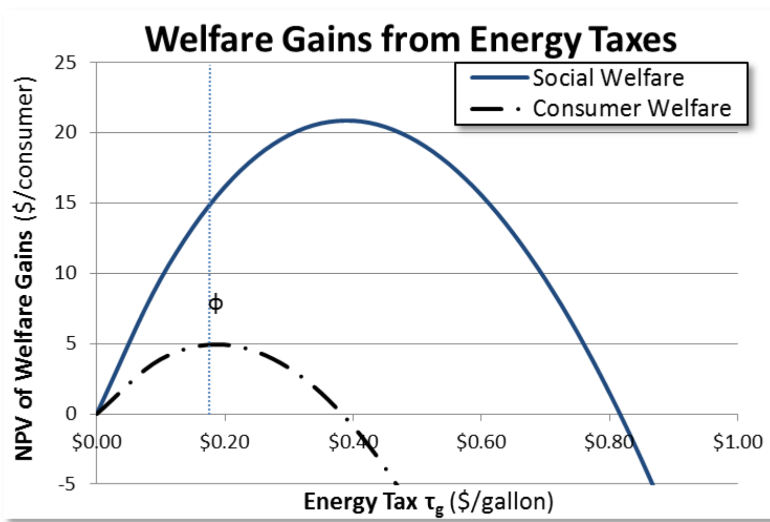
Notes: The dot-dashed green line reflects the demand curve under inattention after the energy tax τ_g is applied. The dotted black line reflects the new supply curve after the product subsidy τ_E is applied.

Figure 3: Heterogeneity and Targeting



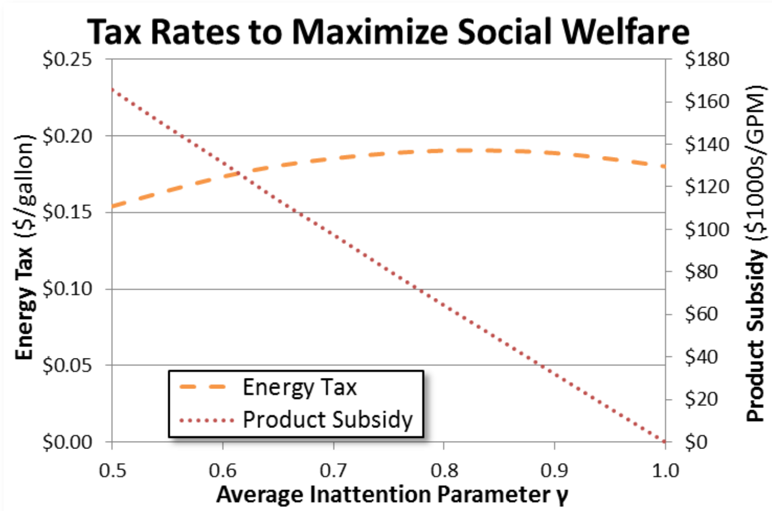
Notes: This figure illustrates the consumer welfare effects of the subsidy when there are two types, rationals and inattentives. The solid blue trapezoid represents allocative gains for the inattentives. The lined red triangle represents allocative losses for the rationals.

Figure 4: The Internality Dividend from Energy Taxes



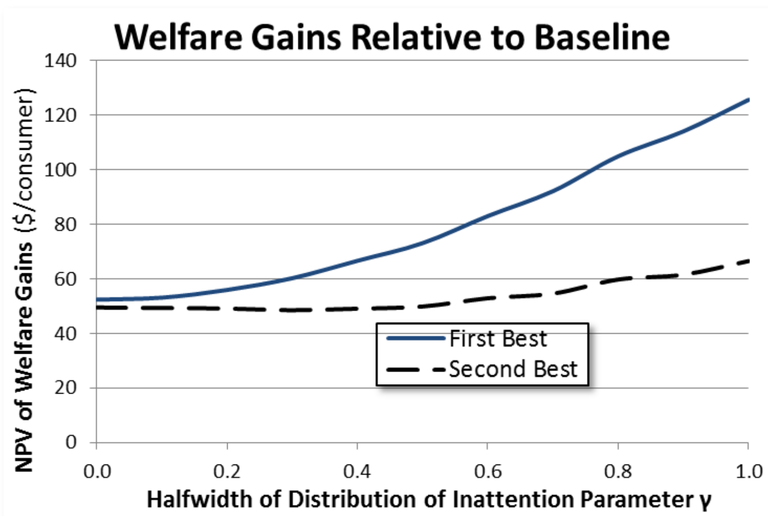
Notes: This figure shows the simulated vehicle market welfare gains from different levels of energy taxes, with the product subsidy set to zero.

Figure 5: Socially-Optimal Policies Under Different Levels of Inattention



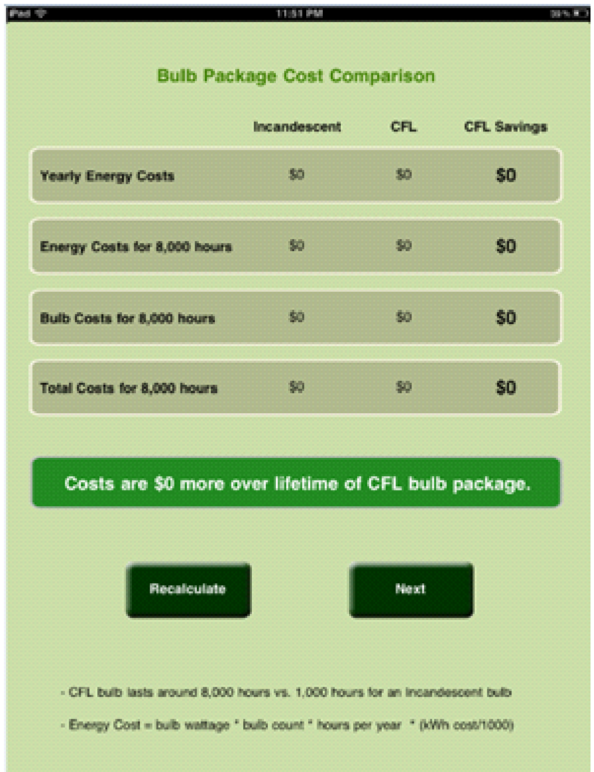
Notes: This figure shows the levels of the energy tax and product subsidy that maximize social welfare as inattention changes.

Figure 6: Implications of Heterogeneous Inattention



Notes: This figure shows that social welfare gains relative to baseline as a function of the degree of heterogeneity in the inattention parameter, which is captured by the halfwidth of the triangular distribution of γ. In all simulations, the peak of the triangular distribution is $\gamma = 0.75$. Second Best refers to the social welfare-maximizing combination of product subsidies and energy taxes. The First Best could be generated by a combination of energy taxes at the level of marginal damages and individual-specific product subsidies.

Figure 7: Lightbulb Information Treatment Screen for iPad



Notes: This screen was shown to the Information Treatment group as part of the lightbulb experiment.

Appendix I: Proofs

Preliminaries

We begin with a series of core results that will be used throughout subsequent proofs.

Lemma 1 *Set* $m^*(\theta, e, p_g) \equiv argmax\{u(m-\theta) - p_g me\}$. *Then* $\frac{\partial}{\partial\theta} m^*(\theta, e, p_g) = 1$.

Proof. Follows from differentiation of the first order condition for m^* and basic algebra. ■

Lemma 2 $em^*(\theta, g, e)$ *is increasing in* e

Proof. We have

$$\frac{\partial}{\partial e} em^*(\theta, g, e) = m^*(\theta, g, e) + e\frac{\partial}{\partial e} m^*(\theta, g, e). \tag{15}$$

Differentiating the first order condition $v'(m^* - \theta) - ge = 0$ with respect to e yields $v''(m^* - \theta)\frac{\partial m^*}{\partial e} = g$. Thus

$$\frac{\partial}{\partial e} em^* = m^* - \frac{eg}{v''(m^* - \theta)} = m^* - \frac{v'(m^* - \theta)}{v''(m^* - \theta)}.$$

But $m - \theta > \frac{v'(m-\theta)}{v''(m-\theta)}$ by assumption, and thus the expression (15) is positive. ■

Lemma 3 $v(\theta, e_E, p_g) - v(\theta, e_I, p_g)$ *is increasing in* θ *and* p_g.

Proof. From the envelope theorem and lemma 1, we have that

$$\frac{\partial}{\partial\theta} v(\theta, e, p_g) = -p_g e.$$

Thus

$$\frac{\partial}{\partial\theta}[v(\theta, e_E, p_g) - v(\theta, e_I, p_g)] = p_g(e_I - e_E) > 0. \tag{16}$$

We also have

$$\frac{\partial}{\partial p_g} v(\theta, e, p_g) = -m^* e$$

and thus

$$\frac{\partial}{\partial p_g}[v(\theta, e_E, p_g) - v(\theta, e_I, p_g)] = m^*(\theta, e_I, p_g)e_I - m^*(\theta, e_E, p_g)e_E \tag{17}$$

■

The next lemma is the key to many of the results in the paper, as it characterizes the marginal consumer between E and I as a function of the various policies and attention. To begin, define the perceived benefit to purchasing E over I for consumers with attention weight γ to be

$$B(\theta, \gamma, p_g, \Delta p) \equiv \gamma p_g[e_I m^*(\theta, e_I, p_g) - e_E m^*(\theta, e_E, p_g)] + \gamma[u(m^*(\theta, e_E, p_g) - \theta) - u(m^*(\theta, e_I, p_g) - \theta)] - \Delta p$$

where $\Delta p = p_E - p_I$ is the difference in prices. Notice that by Lemma 1, however, $u(m^*(\theta, e_E, p_g) - \theta) - u(m^*(\theta, e_I, p_g) - \theta)$ is constant over all θ, and thus we can rewrite as a function $\Delta u(e_E, e_I, p_g) \equiv u(m^*(\theta, e_E, p_g) - \theta) - u(m^*(\theta, e_I, p_g) - \theta)$. Define also $\theta^\dagger$ to satisfy $B(\theta^\dagger, k, p_g, \Delta p) = 0$. Then we have:

Lemma 4 *1.* $\frac{\partial}{\partial\gamma}\theta^\dagger = -\frac{\Delta p}{\gamma p_g(e_I - e_E)} < 0$

2. $\frac{\partial}{\partial \Delta p}\theta^{\dagger} = \frac{1}{\gamma(p_g(e_I - e_E))} > 0$

3. $\frac{\partial}{\partial p_g}\theta^{\dagger} = -\frac{\Delta p - \Delta u(e_E, e_I, p_g)}{\gamma(p_g(e_I - e_E))} < 0$

Proof. First, note that $\frac{\partial}{\partial \theta}B = p_g(e_I - e_E)$, as shown in equation (16).

To prove 1, notice that $\frac{\partial}{\partial \gamma}B = v(\theta, e_E, p_g) - v(\theta, e_I, p_g)$, and since $v(\theta^{\dagger}, e_E, p_g) - v(\theta^{\dagger}, e_I, p_g) = \Delta p/\gamma$ by definition, differentiating $B(\theta_L^{\dagger}, \gamma, p_g, \Delta p) = 0$ with respect to γ using the envelope theorem yields

$$\frac{\partial}{\partial k}\theta^{\dagger} = -\frac{\frac{\partial}{\partial k}B}{\frac{\partial}{\partial \theta}} = -\frac{\Delta p}{\gamma p_g(e_I - e_E)}.$$

Part 2 is proven likewise, using the envelope theorem. Note that $\frac{\partial}{\partial \Delta p}B = -1$, and then differentiate $B(\theta_L^{\dagger}, k, p_g, \Delta p) = 0$ with respect to Δp.

Part 3 is proven similarly, noting $\Delta p - \Delta u(e_E, e_I, p_g) = m^*(\theta, e_I, p_g)e_I - m^*(\theta, e_E, p_g)e_E > 0$ by Lemma 2. ■

Lemma 5 *The function $M(\theta, e, p_g) \equiv v(\theta, e, p_g) + (p_g - c_g - \phi)em^*(\theta, e, p_g)$ is strictly concave and differentiable in p_g and attains its maximum at $p_g = c_g + \phi$.*

Proof. Since

$$\frac{\partial}{\partial p_g}v(\theta, e, p_g) = m^* e$$

some algebra shows that

$$\frac{\partial}{\partial p_g}M(\theta, e, p_g) = (p_g - c_g - \phi)e\frac{\partial}{\partial p_g}m^*(\theta, e, p_g).$$

Since

$$\frac{\partial}{\partial p_g}m^*(\theta, e, p_g) < 0$$

we know that $\frac{\partial}{\partial p_g}M(\theta, e, p_g)$ is positive for $p_g < c_g + \phi$ and negative for $p_g > c_g + \phi$. ■

Lemma 6 *$M(\theta, e_E, p_g) - M(\theta, e_I, p_g)$ is increasing in θ.*

Proof. Differentiating the quantity with respect to θ and using Lemma 1 and equation (16) yields

$$\frac{\partial}{\partial \theta}[M(\theta, e_E, p_g) - M(\theta, e_I, p_g)] = (c_g + \phi)(e_I p_g - e_E p_g) > 0.$$

■

Proofs of claims and propositions in paper

Proof of Claim 1. Obviously the proposed policy achieves the first best.

We now check that no other policy achieves the first best. First, notice that by Lemma 5, $\tau_g = \phi$ in any policy that achieves the first best; otherwise the intensive margin choice will be inefficient. Now with τ_g fixed at ϕ, notice that $\tau_E \neq 0$ creates an inefficiency in the extensive margin choice of durables. ■

Proof of Proposition 1. A bit more generally, we show that $\tau_g^* > \phi$ if the government maximizes W. This will prove the desired result by setting $\phi = 0$.

Let $\theta_L^\dagger$ and $\theta_H^\dagger$ correspond to the utilization needs of the marginal agents for the two attention types and set $\Delta c \equiv c_E - c_I$. We calculate $\frac{\partial}{\partial \tau_g} W$:

$$\begin{aligned}\frac{\partial W}{\partial \tau_g} &= [M(\theta_L^\dagger, e_E, p_g) - M(\theta_L^\dagger, e_I, p_g) - \Delta c]\frac{\partial \theta_L^\dagger}{\partial \tau_g} f(\theta_L^\dagger) + \left[\int_{\theta \le \theta_L^\dagger} \frac{\partial}{\partial p_g} M(\theta, e_I, p_g) dF(\theta) + \int_{\theta \ge \theta_L^\dagger} \frac{\partial}{\partial p_g} M(\theta, e_E, p_g) dF(\theta)\right] \\ &+ [M(\theta_H^\dagger, e_E, p_g) - M(\theta_H^\dagger, e_I, p_g) - \Delta c]\frac{\partial \theta_H^\dagger}{\partial \tau_g} f(\theta_H^\dagger) + \left[\int_{\theta \le \theta_H^\dagger} \frac{\partial}{\partial p_g} M(\theta, e_I, p_g) dF(\theta) + \int_{\theta \ge \theta_H^\dagger} \frac{\partial}{\partial p_g} M(\theta, e_E, p_g) dF(\theta)\right]\end{aligned}$$

where f is the probability density function of F. The first and second lines of the previous equation correspond to the impact on agents with attentions γ_L and γ_H, respectively. The first bracketed term in each line corresponds to the extensive margin effect, and the second bracketed term in each line corresponds to the intensive margin effect.

By Lemma 5, the intensive margin effect of perturbing τ_g is zero when $\tau_g = \phi$. Next, let θ^* be the type such that in the first best allocation, any type with $\theta > \theta^*$ must purchase E and any type with $\theta < \theta^*$ must purchase I. Now when $\tau_g = \phi$, $M(\theta, e, p_g) = v(\theta, e, p_g)$. Moreover, if $\theta_L^\dagger > \theta^*$ then since $v(\theta^*, e_E, p_g) - v(\theta^*, e_I, p_g) - (c_E - c_I) = 0$ by definition, Lemma 3 implies that $v(\theta_L^\dagger, e_E, p_g) - v(\theta_L^\dagger, e_I, p_g) - (c_E - c_I) > 0$. A similar calculation shows that $v(\theta_H^\dagger, e_E, p_g) - v(\theta_H^\dagger, e_I, p_g) - (c_E - c_I) \ge 0$ if $\theta_H^\dagger \ge \theta^*$. Combining this with Lemma 4 then implies that $\frac{\partial}{\partial \tau_g} W > 0$ whenever $\tau_E = 0$ and $\tau_g = \phi$.

Last, we show that $\tau_E = 0$ and $\tau_g < \phi$ can not constitute an optimal tax policy. Suppose that $\tau_g < \phi$ and $\tau_E = 0$. Consider a consumer with utilization need θ. If this consumer sees a benefit of B to purchasing E, then the benefit to this consumer of purchasing E is at least

$$B/\gamma + (\tau_g - \phi)[e_E m^*(\theta, e_E, p_g) - e_I m^*(\theta, e_I, p_g) > B.$$

The inequality follows from the assumption that $\tau_g < \phi$ and because $e_E m^*(\theta, e_E, p_g) < e_I m^*(\theta, e_I, p_g)$ by Lemma 2. Thus under the proposed tax policy, it is socially optimal for any consumer who is indifferent between E and I to purchase E. By Lemma 4 the marginal impact of increasing τ_g has a positive extensive margin effect. And increasing τ_g when it is negative will also have a positive intensive margin effect. ■

Proof of Proposition 2. We have already shown that $\frac{\partial}{\partial \tau_g} W > 0$. The proof that $\frac{\partial}{\partial \tau_p} W > 0$ follows similarly.

We next show that we must have $\tau_E^* > 0$. Lemma 4 shows that

$$\frac{\frac{\partial}{\partial \tau_E}\theta_L^\dagger}{\frac{\partial}{\partial \tau_E}\theta_H^\dagger} = \frac{\gamma_H}{\gamma_L}.$$

Likewise, Lemma 4 shows that

$$\frac{\frac{\partial}{\partial p_g}\theta_L^\dagger}{\frac{\partial}{\partial p_g}\theta_H^\dagger} = \frac{c_E - c_I - \gamma_L \Delta u(e_E, e_I, p_g)}{c_E - c_I - \gamma_H \Delta u(e_E, e_I, p_g)}.$$

Now since $u'(m^*(\theta^\dagger, e, p_g) - \theta^\dagger) = pe$, we know that the slope between $u(m^*(\theta^\dagger, e_E, p_g) - \theta^\dagger)$ and $u(m^*(\theta_L^\dagger, e_E, p_g) - \theta^\dagger)$ is never greater than pe_I, and thus that

$$\Delta u(e_E, e_I, p_g) < e_I(m^*(\theta^\dagger, e_E, p_g) - m^*(\theta^\dagger, e_I, p_g)) < e_E m^*(\theta^\dagger, e_E, p_g) - e_I m^*(\theta^\dagger, e_I, p_g).$$

Thus $\Delta u(e_E, e_I, p_g) < (c_E - c_I)/2$. Moreover, because $\frac{c_E - c_I - \gamma_L \Delta u(e_E, e_I, p_g)}{c_E - c_I - \gamma_H \Delta u(e_E, e_I, p_g)}$ is increasing in Δu, this quantity is maximized at $\Delta u = (c_E - c_I)/2$ and thus

$$\frac{c_E - c_I - \gamma_L \Delta u}{c_E - c_I - \gamma_H \Delta u} < \frac{2-\gamma_L}{2-\gamma_H}$$

Because the function $x(2-x)$ is increasing on $[0,1]$, we also have that $\gamma_L(2-\gamma_L) < \gamma_H(2-\gamma_H)$ and so

$$\frac{2-\gamma_L}{2-\gamma_H} < \frac{\gamma_H}{\gamma_L}.$$

This means that increasing τ_E by a small amount has a larger relative effect on the extensive margin choice of the γ_L agents than does increasing the energy tax by a small amount. Thus $\tau_E = 0$, $\tau_g > \phi$ can't be an optimal combination since an increase in τ_E and an appropriate decrease in τ_g can both reduce the intensive margin distortion while improving the extensive margin choice. ■

Proof of Proposition 3. As will be shown in Proposition 7, W^{SB} is the same under G and G'. Thus it only needs to be shown that if W^{SB} is achieved under the policy $\tau^* = (\tau_E^*, \tau_g^*)$ when the distribution is G, then W^{SB} is achieved by $\tau^{**} = (\tau_E^{**}, \tau_g^*)$ when the distribution is G'. To do this, just check that if $\theta_L^\dagger$ and $\theta_H^\dagger$ are the utilization needs of the marginal γ_L and γ_H consumers under τ^* and G, then $\theta_L^\dagger$ and $\theta_H^\dagger$ will also be the utilization needs of the marginal γ_L and γ_H consumers under τ^{**}. ■

Proof of Proposition 4. Let G be a distribution of attention with weights γ_H and γ_L, and let $G(k)$ be a "scaled down" version of G with weights $k\gamma_H$ and $k\gamma_L$. Let $W_{energy}^{TB}(k)$ be the third-best welfare corresponding to $G(k)$. We will show that $W_{energy}^{TB}(k)$ is decreasing in k. This will complete teh proof as Proposition 7 shows that W^{SB} is constant in k.

Lemma 4 shows that

$$\frac{\frac{\partial}{\partial \gamma}\theta_L^\dagger}{\frac{\partial}{\partial \gamma}\theta_H^\dagger} = \frac{\gamma_H}{\gamma_L}.$$

and as shown in Proposition 2,

$$\frac{\frac{\partial}{\partial p_g}\theta_L^\dagger}{\frac{\partial}{\partial p_g}\theta_H^\dagger} = -\frac{c_E - c_I - \gamma_L \Delta u(e_E, e_I, p_g)}{c_E - c_I - \gamma_H \Delta u(e_E, e_I, p_g)}.$$

Now the optimal energy tax satisfies $\tau_E^* > \phi$ by the proof in Proposition 1. As in the proof of Proposition 1, we decompose the effects of increasing τ_g into the extensive margin effect and the intensive margin effect. By Lemma 5, the intensive margin effect is negative, and thus the assumption that τ_g is set optimally implies that the extensive margin effect is positive. It is also easy to verify that the energy tax must be set such that the γ_L agents are still underpurchasing E (relative to the social optimum) while the γ_H agents are overpurchasing E. Thus increasing k by a small amount has a net positive effect on social welfare because it has a positive extensive margin effect and a zero intensive margin effect. ■

Proof of Proposition 5. Let θ^* be the type such that in the first best allocation, any type with $\theta > \theta^*$ must purchase E and any type with $\theta < \theta^*$ must purchase I.

For this type, the *perceived* gain from purchasing E is

$$\gamma[v(\theta^*, e_I, p_g) - v(\theta^*, e_E, p_g)] - (p_E - p_I).$$

Thus this type will be indifferent between E and I when $\tau_g = \phi$ if and only if

$$\tau_E = (1-\gamma)[v(\theta^*, e_I, p_g) - v(\theta^*, e_E, p_g)].$$

By Lemma 3, a consumer will purchase E if and only if $\theta > \theta^*$, and thus the consumer choice will be first best at this policy. An argument analogous to the proof of Claim 1 shows that no other policy achieves the first best. ■

Proof of Proposition 6, part 1. Let $W_L(\tau)$ and $W_H(\tau)$ correspond to the social welfare of each attention group γ_L and γ_H, so that $W(\tau) = \alpha W_L(\tau) + (1-\alpha)W_H(\tau)$. Now keep γ_L and γ_H constant, and let A be the set of α such that there is an optimal tax policy τ^* under which $W_L(\tau^*) > W_H(\tau^*)$. Set $\alpha^\dagger = \sup A$. First, we claim that W^{SB} is decreasing in α for $\alpha < \alpha^\dagger$. This follows simply because $\alpha W_L + (1-\alpha)W_H$ is decreasing in α when $W_L > W_H$, and thus if $\alpha_1 < \alpha_2 < \alpha^\dagger$, then any second best welfare level achievable under α_2 is also achievable under α_1. Analogous logic shows that W^{SB} is increasing in α when $\alpha > \alpha^\dagger$. ■

Proof of Proposition 6, part 2. Set $r = \gamma_H/\gamma_L$ and assume that $k > 1$. Let $\theta_L^\dagger(r)$ and $\theta_H^\dagger(r)$ be the utilization needs corresponding to the γ_L and γ_H consumers that are on the margin when a second best tax policy $\tau^* = (\tau_E^*, \tau_g^*)$ is implemented. Now consider a different distribution of attention G' in which $\gamma'_H/\gamma'_L = r' < r$, and consider a tax policy (τ_E^{**}, τ_g^*) such that the utilization need of the γ_H consumer who is indifferent between E and I under this policy is still $\theta_H^\dagger(r)$. We will be done if we can just show that under τ^{**} and attention H', the utilization need $\theta_L^\dagger(\tau_E^{**})$ of the marginal γ_L consumer is lower than $\theta_L^\dagger(r)$. To see why this is enough to complete the proof, let θ^{**} be the utilization type such that if the energy tax is set at $\tau_g = \tau_g^*$, then it is socially optimal for any type $\theta < \theta^{**}$ to purchase I and socially optimal for any type $\theta > \theta^{**}$ to purchase E. Such threshold type θ^{**} exists by Lemma 6. But standard envelope theorem arguments imply that τ_E^* is such that $\theta_H^\dagger(r) < \theta^{**} < \theta_L^\dagger(k)$. Thus if $\theta_L^\dagger(\tau^{**}) > \theta^{**}$ we are done, since that implies that under τ^{**} and distribution G', the choices of the more attentive consumers are the same, while the choices of the less attentive consumers are more efficient.

On the other hand, if $\theta_L^\dagger(\tau^{**}) < \theta^{**}$ then we can increase τ_E^{**} to a level τ_E^{***} such that the utilization demand of of the marginal γ_L consumer equals θ^{**} while the utilization demand of the marginal γ_H consumer is now higher than $\theta_H^\dagger(k)$. Then choices of the less and more attentive consumers are again more efficient under G' and the adjusted tax policy.

To finish, some algebra shows that if $\theta_L^\dagger$ and $\theta_H^\dagger$ are the marginal utilization needs under some tax policy, then

$$v(\theta_L^\dagger, e_E, p_g) - v(\theta_L^\dagger, e_I, p_g) = \frac{\gamma_H}{\gamma_L}(v(\theta_H^\dagger, e_E, p_g) - v(\theta_H^\dagger, e_I, p_g)).$$

Thus by Lemma 3, if $\theta_H^\dagger$ is held constant while γ_H/γ_L decreases, $\theta_L^\dagger$ will decrease. ■

Appendix II: Vehicle Market Simulation Details

For the vehicle market simulation, we assume a CRRA functional form for $u(m-\theta)$:

$$u(m_{ij} - \theta_i) = \frac{A}{1-r}(m_{ij} - \theta_i)^{1-r} \tag{18}$$

Given this functional form, the choice of m_{ij} that maximizes utility in Equation (12) is:

$$m_{ij}^* = \theta_i + \left(\frac{\eta p_g e_j}{A}\right)^{-1/r} \quad (19)$$

The parameter r is related to the price elasticity of demand, which we assumed to be -0.15. Specifically, if $\eta < 0$ is the absolute value of elasticity of demand for vehicle-miles traveled (VMT) with respect to p_g, r is:

$$r = \frac{1}{-\eta} \frac{m_{ij}^* - \theta_i}{m_{ij}^*} \quad (20)$$

We assume a uniform distribution of θ, with support ranging from zero to twice the mean. We set A such that $\overline{\theta} = \frac{\overline{m}}{2}$, which ensures that elasticity does not vary too much over the support of θ.

The mean value of θ is set to match nationally-representative data on VMT from the 2001 National Household Travel Survey (NHTS), the most recent national survey with available odometer readings. As part of the survey, odometer readings for about 25,000 vehicles were recorded twice, with several months between the readings, and these data were used to estimate annualized VMT. The variable θ captures potential VMT if the vehicle lasts all the way through an assumed 25-year maximum lifetime. To calibrate θ, we use the NHTS data to calculate the nationally-representative average annual VMT $\overline{m}_a^*$ for vehicles of each age a from 1 to 25. For example, these average annual VMTs decline from 14,500 when new to 9,600 at age 12 and 4,300 at age 25. The U.S. average VMT over a 25-year potential lifetime is the sum of annualized VMT at each age:

$$m^* = \sum_{a=1}^{25} \overline{m}_a^* \approx 236,000 \quad (21)$$

We must translate this undiscounted sum over a potential lifetime to a discounted sum over an expected lifetime. To do this, we apply a scaling factor Λ. We assume a six percent discount rate, which reflects the average discount rate for vehicle buyers calculated by Allcott and Wozny (2011), giving a discount factor $\delta = \frac{1}{1.06}$. We use data on nationwide registrations of new and used vehicles from 1999 to 2008 to calculate the average survival probability of vehicles at each age a. These are multiplied to construct cumulative survival probabilities, denoted ϕ_a. For example, a new vehicle has a 60 percent chance of surviving to age 12 and a ten percent chance of surviving to age 25. The scaling factor Λ is:

$$\Lambda = \frac{\sum_{a=1}^{25} \delta^a \overline{m}_a^* \phi_a}{\sum_{a=1}^{25} \overline{m}_a^*} \approx 0.436 \quad (22)$$