The Lightbulb Paradox: Evidence from Two Randomized Experiments

Hunt Allcott and Dmitry Taubinsky*

November 26, 2013

Abstract

It is often suggested that consumers are imperfectly informed about or inattentive to energy costs of durable goods such as cars, air conditioners, and lightbulbs. We study two randomized control experiments that provide information on energy costs and product lifetimes for energy efficient compact fluorescent lightbulbs (CFLs) vs. traditional incandescent bulbs. We then propose a general model of consumer bias in choices between energy-using durables, derive sufficient statistics for quantifying the welfare implications of such bias, and evaluate energy efficiency subsidies and standards as second best corrective policies if powerful information disclosure is infeasible. In the context of our theoretical model, the empirical results suggest that moderate CFL subsidies may be optimal, but imperfect information and inattention do not appear to justify a ban on traditional incandescent lightbulbs in the absence of other inefficiencies.

JEL Codes: D03, D12, H21, H31, L94, Q41, Q48.

Keywords: Behavioral public economics, energy efficiency, energy-using durables, field experiments, internality taxes, optimal taxation.

We are grateful to Lucas Davis, Stefano DellaVigna, Mushfiq Mubarak, Sendhil Mullainathan, Emmanuel Saez, and other colleagues, as well as seminar audiences at Berkeley, Harvard, the NBER Public Economics Meetings, Stanford, the Stanford Institute for Theoretical Economics, and the University of California Energy Institute for constructive feedback. We thank our research assistants - Jeremiah Hair, Nina Yang, and Tiffany Yee - as well as management at the partner company, for their work on the in-store experiment. Thanks to Stefan Subias, Benjamin DiPaola, and Poom Nukulkij at GfK for their work on the TESS experiment. We are grateful to the National Science Foundation and the Sloan Foundation for financial support. Code to replicate the results, as well as the protocol for the in-store experiment, can be downloaded from Hunt Allcott's website.

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

1 Introduction

It has long been suggested that consumers may be imperfectly informed about or inattentive to energy costs when they buy energy using durables such as cars, air conditioners, and lightbulbs.[1] This suggestion is supported by recent empirical evidence from other domains: people are inattentive to "add-ons" or ancillary product costs such as sales taxes (Chetty, Looney, and Kroft 2009),[2] shipping and handling charges (Hossein and Morgan 2006), and out-of-pocket insurance costs (Abaluck and Gruber 2011). Because American households spent $325 billion on gasoline and another $245 billion on electricity, natural gas, and heating oil in 2011 (BLS 2013), even small inefficiencies can aggregate to substantial losses.

In theory, the first best policy to address imperfect information and inattention would be an idealized information provision technology that is both costless and "powerful", by which we mean that all treated consumers would become fully informed about and attentive to energy costs and other product attributes. In practice, the U.S. and other countries have energy use disclosure requirements such as fuel economy labels on new cars and "yellow tags" on home appliances. There is little evidence on how effectively these programs inform consumers or how they affect purchases.

In addition to information disclosure, policymakers also have a broad set of second best corrective policies such as subsidies and standards for energy efficient autos, appliances, and buildings. Along with externalities and the so-called "landlord-tenant" agency problem, imperfect information and inattention are key potential justifications in both academic papers[3] and government regulatory impact analyses.[4] Evaluating these policies is important, as they are costly: fuel economy standards, appliance energy efficiency standards, "demand-side management" programs run by electric and gas utilities, and weatherization subsidies cost $17 billion each year (Allcott and Greenstone 2012).

This paper focuses on one the lightbulb market, a particularly compelling case study of consumer behavior and public policy. Do people know how much electricity a traditional incandescent uses relative to an energy efficient compact fluorescent lightbulb (CFL)? Do we pay as much attention to this additional cost as we do to purchase prices, or are we inattentive, like consumers in Gabaix and Laibson (2006)? How should the government intervene, if at all? Regulated and

[1] See Anderson and Claxton (1982), Blumstein *et al.* (1980), Jaffe and Stavins (1994), Sanstad and Howarth (1994), Gillingham and Palmer (2013), and many others.

[2] Relatedly, Finkelstein (2009) and Cabral and Hoxby (2012) show how salience varies across different tax collection methods.

[3] Among other references, see Gillingham and Palmer (2013), Fischer, Harrington, and Parry (2007), and Parry, Evans, and Oates (2010). The latter paper, for example, focuses on two market failures that could justify energy efficiency standards: externalities and what they call "misperceptions market failures."

[4] See, for example, the Regulatory Impact Analysis for the increase in the Corporate Average Fuel Economy standard for 2012 to 2016. The analysis argues that even without counting the externality reductions, the regulation increases consumer welfare, perhaps because consumers have incorrect "perceptions" of the value of fuel economy (NHTSA 2010, page 2). See also the Regulatory Impact Statement for Australia's ban on energy inefficient lightbulbs (DEWHA 2008, page vii), which argues that "information failures" and consumer cognitive costs help to justify that policy.

government-run electric utilities spent $252 million subsidizing and otherwise promoting energy efficient compact fluorescent lightbulbs (CFLs) in the U.S. in 2010 (DOE 2010). Furthermore, the Energy Independence and Security Act of 2007 sets minimum efficiency standards that ban traditional incandescent lightbulbs between 2012 and 2014 and will be tightened further in 2020.

This ban on traditional incandescents has generated vigorous debate. Many consumers dislike CFLs because they are inferior on several dimensions, and opponents suggest that the regulation is "an example of over-reaching government intrusion into our lives" (Formisano 2008). The government's economic impact analysis (DOE 2009) and other studies (NRDC 2011) argue for the policy by showing that consumers will enjoy billions of dollars in annual cost savings. Of course, such private cost savings could only reflect welfare gains in the presence of imperfect information, inattention, or some other market failure. Argentina, Australia, Brazil, Canada, China, Cuba, the European Union, Israel, Malaysia, Russia, and Switzerland have also banned some or all incandescent light bulbs.

We combine two randomized information provision experiments with a formal model of optimal policy to answer two research questions. First, *how much can information provision affect demand for energy efficient lightbulbs?* Second, *if powerful information provision is costly or infeasible, do subsidies and minimum standards increase welfare as second best solutions to imperfect information and inattention?*

We begin by presenting the results of two randomized control trials (RCTs) that measure the effects of energy cost information on lightbulb purchases. The first is an "artefactual field experiment" (Levitt and List 2009) using Time-Sharing Experiments for the Social Sciences (TESS). This is a high-quality computer-based survey platform which has been used by a number of economists, including Allcott (2013), Fong and Luttmer (2009), Heiss, McFadden, and Winter (2007), Newell and Siikamaki (2013), Rabin and Weizsacker (2009), and others. We gave consumers a $10 shopping budget and asked them to make a series of choices between CFLs and incandescents in a multiple price list format. We then gave the treatment group information about lightbulb energy costs and lifetimes. After this informational intervention, we again asked all consumers to choose between CFLs and incandescents. This design allows us to infer the joint distribution of demand in the baseline and "informed" states, which is crucial for policy analysis. The experiment was incentive compatible: one of the choices was randomly selected to be the consumer's "official purchase," and the consumer received those lightbulbs and kept the remainder of his or her shopping budget. The informational intervention increased average willingness to pay for CFLs by $2.32, and CFL market share at market prices increased by about 12 percentage points.

The second RCT is a "framed field experiment" (Levitt and List 2009) with a large home improvement retailer. Our staff intercepted shoppers, used an iPad to deliver information about energy costs and bulb lifetimes to the treatment group, and then gave coupons with randomly assigned CFL subsidies. While a 20 percent subsidy increased CFL market share by about 10 percentage points, the informational intervention had statistically zero effect. We can bound the information effect as less than the effect of a 12 percent subsidy with 90 percent confidence.

The second half of the paper uses the experimental results to analyze the welfare effects of second best subsidies and standards, given that our informational interventions are not realistically feasible at large scale. In doing this, we follow the approach of Chetty, Looney, and Kroft (2009) in assuming that our treatment groups made informed and otherwise optimal decisions, meaning that our treatment effects measure the magnitude of bias from imperfect information and inattention. Qualitatively, we believe that this assumption is best viewed as an approximation, and we present evidence to evaluate it throughout the paper. In order to ensure that the assumption would be particularly reasonable, we designed the two interventions to fully inform consumers and draw attention to energy costs and other attributes, while minimizing other possible effects. For example, the interventions included no information or cues related to environmental externalities, other social costs of energy use, or social norms. We also took a series of steps to minimize experimenter demand effects. Furthermore, we delivered information through different channels and quizzed consumers on comprehension in the TESS experiment, thus ensuring that consumers understood and attended to the information.

In order to use the empirical results for policy analysis, we formalize a simple theoretical framework that clarifies the "internality rationale" for energy efficiency policy. Consumers make a discrete choice, which in our application is between an incandescent and a CFL. Some consumers, however, may misjudge the true difference in utility they would experience from the two products; we label the dollar value of this potential mistake the "internality." The internality is directly analogous to an externality: it is a wedge between willingness to pay and social welfare, and it may be heterogeneous across consumers. The policymaker has two instruments: an "internality tax" (in our example, a CFL subsidy) with lump-sum recycling and a ban on the "sin good" (in our example, a ban on traditional incandescents). Just as Diamond (1973) shows that the optimal externality tax equals the average marginal externality, the optimal internality tax equals the average marginal internality. The welfare effect of the ban is the sum of the true utility experienced by the set of consumers who would buy the banned good if they were allowed to do so. Crucially, this average marginal internality is a sufficient statistic in the sense of Chetty (2009): the underlying "structural" model of the bias and any heterogeneity within the set of marginal consumers are both irrelevant for evaluating the welfare effects of a subsidy or ban.

In the context of our theoretical model, the TESS experiment results suggest that the optimal subsidy is approximately $3 per 60-Watt equivalent CFL. This is slightly larger than typical CFL subsidies offered by many electric utilities. However, we also observe a large group of consumers who purchase incandescents at baseline and are still willing to pay substantially more for incandescents after the informational intervention. Banning incandescents imposes welfare losses on this population that outweigh the gains to apparently-biased consumers who had weaker preferences for the incandescent. This implies that in our model, imperfect information and inattention by themselves do not justify a ban on traditional incandescents.

The simpler design of the in-store experiment identifies only the slope of demand and the effect of information on quantity demanded. However, we show how these two parameters can be used to derive a first-order approximation to the optimal subsidy. Intuitively, the average marginal internality is the price change that would have the same effect on demand as the informational intervention. Given that the intervention had statistically zero effect, we cannot reject that the optimal subsidy is zero. Our formula bounds the optimal subsidy for a 60-Watt equivalent CFL between negative 30 cents and positive 35 cents per CFL with 90 percent confidence. Given the difference in the population and the experimental setting, it does not surprise us that the effects differ from the TESS experiment. This result only strengthens the qualitative conclusion that the internalities we consider are not large enough to solely justify a ban in our model.

The paper makes three central contributions. First, our two experiments are a "proof of concept" for how large-sample randomized control trials can be used to test the effects of energy use information on durable goods purchases.[5] The dearth of evidence in this context is especially remarkable given the large literature on the effects of information disclosure on consumer choice in other domains, including Choi, Laibson, and Madrian (2010) and Duarte and Hastings (2012) on financial choices, Greenstone, Oyer, and Vissing-Jorgensen (2006) on securities, Bhargava and Manoli (2013) on takeup of social programs, Jin and Sorensen (2006), Kling *et al.* (2012), and Scanlon *et al.* (2002) on health insurance plans, Pope (2009) on hospitals, Bollinger, Leslie, and Sorensen (2011) and Luo *et al.* (2012) on health and nutrition, Dupas (2011) on HIV risk, Figlio and Lucas (2004) and Hastings and Weinstein (2008) on school choice, and many others.

Second, there is a growing empirical literature on whether consumers of durable goods "undervalue" energy costs relative to upfront prices, including Allcott (2013), Allcott and Wozny (2013), Busse, Knittel, and Zettelmeyer (2013), Dubin and McFadden (1984), Goldberg (1998), Hassett and Metcalf (1995), Hausman (1979), Metcalf and Hassett (1999), Sallee, West, and Fan (2009), and many others. Imperfect information and inattention are two of the factors that could cause undervaluation. Most of the previous literature has tested for undervaluation by (essentially) comparing price elasticities to energy cost elasticities using variation in purchase prices and energy costs. Our approach is innovative in this literature in that we instead test for undervaluation using experimentally-induced variation in information about and salience of energy costs, as suggested by Chetty, Looney, and Kroft (2009) and DellaVigna (2009). Aside from allowing for more highly-credible identification via randomized control experiments, this approach also isolates the potential

[5] There are some related studies that differ from our experiments on one or more dimensions. Kallbekken, Saelen, and Hermansen (2013) study energy information disclosure at six retail stores in Norway, comparing purchases to a non-randomly selected control group. Anderson and Claxton (1982) study energy information labels with 12 stores assigned to treatment groups and six to control. There are a number of studies that randomly assign information disclosure across individual experimental subjects and study effects on stated preferences in hypothetical choices, including Newell and Siikamaki (2013) and Ward, Clark, Jensen, Yen, and Russell (2011). Deutsch (2010a, 2010b) studies information disclosure to online shoppers, measuring what products they click on and what products they put in online shopping carts, but he does not observe actual purchases. Houde (2012) uses quasi-experimental variation with a structural demand model to estimate how the Energy Star label affects consumer welfare, while Herberich, List, and Price (2011) and Toledo (2013) study how prices and social norm information affect CFL purchases.

effects of imperfect information and inattention from other potential mechanisms such as present bias, which should be unaffected by our informational interventions. Our results are qualitatively consistent with several of the above papers in suggesting that internalities are small in the particular markets that have been studied, and that corrective subsidies and standards may be stronger in these contexts than can be justified by internalities alone.

Third, we provide an example of how techniques from public economics can be combined with psychologically-motivated experiments to provide insight into important public policies. Related analyses include Bernheim and Rangel (2004), Chetty, Looney, and Kroft (2009), Gruber and Koszegi (2004), Gul and Pesendorfer (2007), and O'Donoghue and Rabin (2006), who study taxes when consumers are present biased or otherwise make mistakes. There are also several analyses of energy taxes, energy efficiency standards, or subsidies for energy efficient goods when consumers misoptimize, including Allcott, Mullainathan, and Taubinsky (2013), Heutel (2011), Fischer, Harrington, and Parry, and Parry Evans, and Oates (2010). Our theoretical framework is relatively straightforward, and it is closely related to Allcott, Mullainathan, and Taubinsky (2013). This paper is distinguished from most existing work in "behavioral public economics" in that it combines a theoretical framework with parameters from randomized experiments to derive optimal policy.

Section 2 gives more background on lightbulbs and related policies. Sections 3 and 4 present the TESS and in-store experiments, respectively. Section 5 lays out our theoretical framework and derives optimal policies and welfare formulas. Section 6 contains the policy evaluations, and Section 7 concludes.

2 Background

2.1 "The Lightbulb Paradox"

Lightbulbs are a canonical example of the "Energy Paradox" (Jaffe and Stavins 1994): the low adoption of energy efficient technologies despite potentially large savings. Compared to standard incandescents, compact fluorescent lightbulbs (CFLs) typically last eight times longer and use four times less electricity. Although CFLs cost several dollars to purchase, compared to a dollar or less for incandescents, using a CFL saves about \$5 each year once the costs of electricity and replacement bulbs are included. Despite this cost advantage, only 28 percent of residential sockets that could hold CFLs in 2010 actually had them (DOE 2010). In that year, using incandescents instead of CFLs cost US households \$15 billion.[6] Although one lightbulb is inexpensive and by itself uses little electricity, this aggregate figure makes it difficult to argue that the lightbulb market

[6] Throughout the paper, we assume that incandescents and CFLs last an average of 1000 and 8000 hours, respectively. (To receive the Energy Star rating, a CFL model must last a median of 8000 hours in official tests. Of course, a given consumer may experience varying results.) The national average electricity price is \$0.10 per kilowatt-hour. Our cost estimate of \$15 billion is equal to 5.8 billion residential sockets (DOE 2012), times the 80 percent of sockets that can accommodate CFLs (DOE 2010) minus the actual "socket share" of 28 percent (DOE 2010), times \$5 per socket per year.

is unimportant, especially when viewed as a case study of issues relevant to the broader class of energy-using durables.

Of course, CFLs and incandescents are differentiated products: many consumers do not like CFLs because the light quality is different, they sometimes flicker, they take time to reach full brightness, and they must be properly disposed of because they contain mercury. While about 60 percent of Americans report in recent surveys that they are "excited" about the lightbulb efficiency standards, about 30 percent say that they are "worried" because they "prefer using traditional lightbulbs" (Sylvania 2012). As Jaffe and Stavins (1994), Allcott and Greenstone (2012), and many others have pointed out, these kinds of non-financial utility costs from energy efficiency are important potential explanations for the apparent "Energy Paradox." Our framework is very clear in allowing these utility differences, and our results indeed show that many consumers strongly prefer incandescents even after the informational interventions.

The U.S. lighting efficiency standards do not require CFLs, nor do they ban incandescents. Instead, they set a maximum energy use per unit of light output. Along with CFLs, light-emitting diodes (LEDs) and high-efficiency halogen bulbs also comply with the standard. We focus on the choice between CFLs and incandescents because these are *by far* the most important current technologies. In 2012, about 1.5 billion incandescents and 300 million CFLs were purchased, compared to only 23 million LEDs (Energy Star 2013). Our quantitative welfare calculations would change in the future if LEDs become a relevant part of the choice set. However, it seems plausible that the qualitative lessons about imperfect information and inattention from CFLs also apply to LEDs, given that LEDs also have high purchase prices, long lifetimes, and large energy cost savings relative to both incandescents and CFLs.

2.2 Economic Reasons for Standards and Subsidies

Governments often intervene in markets to subsidize goods or ban bads. Review articles such as Allcott and Greenstone (2012), Gillingham and Palmer (2013), Jaffe and Stavins (1994), and many others discuss the economic reasons in the context of energy-using durables. One potential reason for such policies is externalities. In the case of lightbulbs, one might think that electricity prices are below social cost due to unpriced externalities from climate change, and banning energy inefficient lightbulbs is a welfare-improving second best policy in the absence of a price on carbon dioxide emissions. However, two other distortions that imply that the marginal price of electricity used for residential lightbulbs could actually be *above* social marginal cost. First, retailers typically include much of fixed distribution costs in marginal prices, as Borenstein and Davis (2012) and Davis and Muehlegger (2010) show for natural gas. Second, most residential customers are charged time-invariant prices instead of the optimal peak-load prices, which are lower at night and higher during the day. If lightbulbs are more likely than to be used at night, they thus use electricity which is underpriced. This suggests that if the primary distortion is mispriced residential electricity, it could actually be optimal to *subsidize* incandescents.[7]

Policymakers might also subsidize new or emerging products to help correct for uninternalized spillovers from research and development or consumer learning. However, 70 percent of consumers report having at least one CFL in their home, compared to 80 percent who report having at least one incandescent (Sylvania 2012), so the technology is available and the vast majority of consumers already have experience with it.

Agency problems in real estate markets could also justify subsidies and standards. For example, home buyers and renters cannot costlessly observe energy efficiency, which reduces the incentive of sellers and landlords to install energy efficient capital stock. Empirical studies by Davis (2010) and Gillingham, Harding, and Rapson (2012) provide some evidence of this, and we are able to provide some evidence from the TESS experiment as well. A final set of inefficiencies is "internalities," which we define as choices that don't maximize the decision maker's own welfare. Present bias over cash flows could be one such internality: if consumers underweight future energy cost savings, they would be less energy efficient than their long-run preferences would dictate, and sophisticates would demand commitment devices to make their future selves buy CFLs and hybrid cars. However, as Andreoni and Sprenger (2012) and others have pointed out, agents in most models are present biased over *consumption*, and most consumers have enough liquidity that paying the incremental few dollars for a CFL does not immediately affect consumption.

Our paper focuses on a particular class of internalities, imperfect information and inattention. In the absence of our results, what would empirical estimates from other contexts suggest could be the magnitudes of these biases? Abaluck and Gruber (2011) find that consumers are five times more responsive to insurance plan premiums than to out-of-pocket costs. In their two empirical studies, Chetty, Looney, and Kroft (2009) estimate that consumers are only 35 percent and 6 percent as attentive to sales taxes as they are to product prices. A CFL saves an undiscounted \$36 over its expected life relative to an incandescent. If consumers were (hypothetically) as inattentive to these savings as they are to sales taxes, these estimates suggest that they could undervalue the CFL's energy savings by \$23 to \$34. This dwarfs the typical difference in purchase prices between CFLs and incandescents, and it suggests that our informational interventions could have massive impacts on demand.

In summary, while there are other market failures that could justify subsidies and standards for lightbulbs, we designed this study to focus on imperfect information and inattention because results from other literatures suggested that these two distortions could be large, while other market failures appear to be less relevant.

[7] California is a particularly stark example. Regulations encouraging low-carbon electricity generation mean that the carbon content of electricity consumed there is extremely low relative to other states, so the downward distortion to electricity prices from the lack of a carbon tax is particularly small. Meanwhile, residential electricity tariffs with sharply increasing block prices distort marginal prices upward. Despite the fact that these two forces significantly weaken or reverse the argument that underpriced electricity justifies energy efficiency policies, California has implemented the federal lighting efficiency standards early.

3 TESS Experiment

3.1 Survey Platform

We implemented the artefactual field experiment through Time-Sharing Experiments for the Social Sciences (TESS). TESS, which is funded by the National Science Foundation, facilitates academic access to KnowledgePanel, an online experimental platform managed by a company called GfK. The platform has been used by a number of economists, including Allcott (2013), Fong and Luttmer (2009), Heiss, McFadden, and Winter (2007), Newell and Siikamaki (2013), and Rabin and Weizsacker (2009).

One reason why economists use TESS is the recruitment process, which generates a sample as close as practically possible to being nationally representative on unobservable characteristics and reduces concern about generalizability. Potential KnowledgePanel participants are randomly selected from the U.S. Postal Service Delivery Sequence File and recruited through a series of mailings in English and Spanish, plus telephone-based follow-up when the address can be matched to a phone number. About 10 percent of people who are invited actually consent and complete the demographic profile to become KnowledgePanel participants, and there are now approximately 50,000 active panel members. Unrecruited volunteers are not allowed to opt in. Households without computers are given computers in order to complete the studies.

KnowledgePanel participants take an average of two studies per month, and no more than one per week. Of the KnowledgePanel participants who started our study, some were not qualified to continue because their computer audio did not work, which would have prevented them from hearing the audio part of our treatment and control interventions. Of the qualified participants who began the survey, about 3/4 completed it, giving us a final qualified sample size of 1533. Although we were able to negotiate with GfK to require answers on some of the most important parts of the study, GfK's policy is to usually allow participants to refuse. A handful of participants refuse to answer any given question, so the number of observations will vary slightly across regressions.

GfK provides sampling weights which can allow us to match the US population aged 18 and older on gender, age, ethnicity, education, census region, urban or rural location, and whether the household had internet access before recruitment. All statistics presented in the paper are weighted for national representativeness, although weighting observations equally does not substantively change the results.

3.2 Experimental Design

The study had four parts: baseline lightbulb choices, the informational intervention, endline lightbulb choices, and a post-experiment survey. About 21 percent of treatment group consumers (226 consumers) were randomly assigned to skip the baseline lightbulb choices and begin directly with the informational intervention; these "Endline-Only" consumers are included in all statistics when possible and are the focus of several robustness checks. Appendix 1 contains screen shots from each part of the experiment.

Consumers were first shown an introductory screen with the following text:

In appreciation for your participation in this study, we are giving you a $10 shopping budget. With this money, we will offer you the chance to buy light bulbs. You must make a purchase with this money. Whatever money you have left over after your purchase, you get to keep. This money will be provided to you as cash-equivalent bonus points that will be awarded to your account.

In approximately four to six weeks, GfK will send you the light bulbs you have purchased. Light bulbs are frequently shipped in the mail. There is not much risk of breakage, but if anything does happen, GfK will just ship you a replacement. Even if you don't need light bulbs right now, remember that you can store them and use them in the future.

During the study, we will ask you to make 30 decisions between pairs of light bulbs. There will be a first set of 15decisions, then a break, and then a second set of 15 decisions. After you finish with all 30 decisions in the questionnaire, one of them will be randomly selected as your "official purchase." GfK will ship you the light bulbs that you chose in that official purchase. Since each of your decisions has a chance of being your official purchase, you should think about each decision carefully.

3.2.1 Baseline Lightbulb Choices

After the introductory screen, consumers were then shown two lightbulb packages. One package contained one Philips 60-Watt equivalent Compact Fluorescent Lightbulb. The other contained four Philips 60-Watt incandescent lightbulbs. The two choices were chosen to be as comparable as possible, except for the CFL vs. incandescent technology. Half of respondents were randomly assigned to see the incandescent on the left, labeled as "Choice A," while the other half saw the incandescent on the right, labeled as "Choice B."

Consumers had the option to "click for detailed product information," and about 19 percent did so. This opened a simple "Detailed Product Information" screen, which included the light output in Lumens, a quantitative measure of light color, energy use in Watts, and other information. Both packages typically sell online for about $4, although we did not tell consumers these typical prices.

Lower down on this same screen, consumers were asked to make their baseline lightbulb choices: 15 decisions between the same two packages at different relative prices. Decision Number 1 offered Choice A for free and Choice B for $10. The relative price of Choice A increased monotonically until Decision Number 15, which offered Choice A for $10 and Choice B for free. Consumers spent a median of three minutes and zero seconds to complete these first 15 decisions.

We identify consumers' baseline relative willingness to pay (WTP) for the CFL, denoted v_i^0, using the relative prices at which they switch from preferring CFLs to incandescents. For example, consumers who choose CFLs when both packages cost $4 but choose incandescents when incandescents are one dollar cheaper are assumed to have $v_i^0 = \$0.50$. Eight percent of consumers did not choose monotonically: they chose Choice A at a higher relative price than another decision at which they chose Choice B. These consumers were prompted with the following message: *The Decision Numbers below are organized such that Choice A costs more and more relative to Choice B as you read from top to bottom. Thus, most people will be more likely to purchase Choice A for decisions at the top of the list, and Choice B for decisions at the bottom of the list. Feel free to review your choices and make any changes. Then click NEXT.* After this prompt, 5.3 percent of consumers still chose non-monotonically, and we code their WTP as missing.

Some consumers had "censored" WTPs: they preferred either Choice A or Choice B at all relative prices. These consumers were asked to self-report their WTP. For example, a participant who always preferred Choice A was asked: *Your decisions suggest that you prefer Choice A even when Choice A costs a total of $10 and Choice B is free. If Choice B continued to be free, how much would Choice A need to cost in order for you to switch to Choice B? (This is purely hypothetical - your answer will not affect any of the prices you are offered.)*

At baseline, five percent of consumers preferred the incandescents by more than $10, while 19 percent preferred the CFLs by more than $10. Across all censored consumers, the median absolute value of self-reported relative WTP was $15. The distribution of self-reports is skewed, with about eight percent of consumers preferring one or the other choice by more than $40. Because these are self-reports, we wish to be cautious about using them in the analysis, so we instead assume a mean relative WTP of $15 and -$15 for top-coded and bottom-coded consumers, respectively. We will demonstrate the sensitivity of the results to this assumed mean censored value. While an assumed mean censored value may seem unsatisfying, remember that in the absence of this type of experiment, a demand model used to predict the removal of a product from the choice set would typically assume a logit or otherwise parametric functional form for demand.

In theory, if lightbulbs were perishable and consumers did not immediately need one, they would buy the cheapest package instead of revealing the WTP they would have if they did need one. In practice, lightbulbs are easily stored, and we reminded consumers of this fact in the introductory text. In theory, if it were costless to resell the experimental purchase and replace it with a different purchase outside the experiment, consumers who know that the typical retail prices are approximately equal would always buy the cheaper package. In order to avoid making this salient, the experiment website did not include information about the bulbs' typical retail prices. In practice, it seems unlikely that consumers resold the packages that they received. If non-storability and price arbitrage affected some consumers' choices, this would make the demand curve more elastic and the treatment effects less positive. Empirically, however, we see large shares of consumers with relative WTPs that differ substantially from typical market relative prices.

3.2.2 Informational Intervention

Consumers were randomized into three groups: Balanced treatment, Positive treatment, and control. Endline-Only consumers were randomized between the two treatment groups with equal probability. All other consumers were randomized between all three groups with equal probability.

The information treatments were designed to give clear product information while minimizing the possibility that the information would be perceived as biased. The treatments were also designed to closely parallel each other, to minimize the chance that idiosyncratic factors other than the information content could affect purchases. Each treatment had the following structure:

1. Belief elicitation. This elicited prior beliefs over the information to be presented in the two Information Screens.

2. Introductory Screen

3. First Information Screen. This had text plus an illustratory graph, and the text was also read verbatim via an audio recording. The audio recordings are available as part of the Online Supplementary Materials. At the bottom of the information screen, there was a "quiz" on a key fact.

4. Second Information Screen. This paralleled the First Information Screen.

The central difference between the three conditions was the content of their two Information Screens. The order of these two screens was randomly assigned with equal probability.

We took two steps to make sure that all consumers understood the treatment. First, we used multiple channels to convey information: text, graphical, and audio. This means that people who process information in different ways had a higher chance of internalizing the information. Second, the quiz forced respondents to internalize the information if they had not done so initially.

Balanced Treatment The Balanced treatment Introductory Screen had the following text:

For this next part of the study, you will have the opportunity to learn more about light bulbs. We will focus on the following two issues:

1. *Total Costs*

2. *Disposal and Warm-Up Time*

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

Consumers then advanced to the Total Cost Information Screen and the Disposal and the Warm-Up Information Screen, in randomized order. The Total Cost Information Screen explained that CFLs both last longer and use less electricity and translates these differences into dollar amounts. The bottom line was:

Thus, for eight years of light, the total costs to purchase bulbs and electricity would be:

- *$56 for incandescents: $8 for the bulbs plus $48 for electricity.*

- *\$16 for a CFL: \$4 for the bulbs plus \$12 for electricity.*

The quiz question at the bottom of the screen was: *For eight years of light, how much larger are the total costs (for bulbs plus electricity) for 60-Watt incandescents as compared to their CFL equivalents?* The correct answer could be inferred from the information on the screen: \$56 for incandescents - \$16 for CFLs = \$40. Sixty-four percent of consumers correctly put \$40. Those who did not were prompted: *That is not the correct answer. Please try again.* After this prompt, 73 percent of consumers had typed \$40. The remaining consumers were prompted: *The total costs for eight years of light are \$16 for CFLs and \$56 for incandescents. Therefore, the incandescents cost \$40 more. You may type the number 40 into the answer box.* By this point, 89 percent of consumers had correctly typed \$40. This documents that the vast majority of consumers understood at least some part of the information. Consumers spent a median of two minutes and 12 seconds to read the Total Cost Information Screen and complete the quiz question.

The Disposal and Warm-Up Information Screen was designed to present information about ways in which CFLs may not be preferred to incandescents. It paralleled exactly the Total Cost Information Screen, beginning with belief elicitation, and then continuing to an Information Screen with text of similar length, a graph, and a quiz question at the bottom. The Disposal and Warm-Up Information Screen explained that "*because CFLs contain mercury, it is recommended that they be properly recycled instead of disposed of in regular household trash.*" It also explained that "*after the light switch is turned on, CFLs take longer to warm up than incandescents.*" We included this information to reduce the probability of experimenter demand effects, through which consumers might think that the experimenter wanted them to purchase the CFL, potentially causing their endline choices to differ from true preferences.

Control The control intervention was designed to exactly parallel the treatment interventions, but with information that should not affect relative WTP for CFLs vs. incandescents. One screen presented the number of lightbulbs installed in residential, commercial, and industrial buildings in the United States. The other screen detailed trends in total lightbulb sales between 2000 and 2009.

Positive Treatment The Positive treatment was designed to inform consumers about the benefits of the CFL in terms of lifetime and lower energy costs, without presenting information about ways that the incandescent might be preferred to the CFL. This more closely parallels the intervention in the in-store RCT described in the next section. To implement this while keeping the intervention the same length, we combined the Total Cost Information Screen with a random draw of one of the two control screens.

3.2.3 Endline Lightbulb Choices

The endline lightbulb choice screen was analogous to the baseline screen. Consumers spent a median of one minute and 20 seconds to complete these final 15 decisions. We determine endline relative WTP v_i^1 in the same way as above.

3.3 Data

Column 1 of Table 1 presents descriptive statistics. Liberal is self-reported political ideology, originally on a seven-point scale, normalized to mean zero and standard deviation one, with larger numbers indicating more liberal. Party is self-reported political affiliation, similarly normalized from an original seven-point scale, with larger numbers indicating more strongly Democratic. Environmentalist is the consumer's answer to the question, "Would you describe yourself as an environmentalist?" Conserve Energy is an indicator for whether the consumer reports having taken steps to conserve energy in the past twelve months. Homeowner is a binary indicator variable for whether the consumer owns his or her home instead of rents. These questions were asked when the participant first entered KnowledgePanel, not as part of our experiment.

Column 2 presents the difference in means between consumers in either of the two treatment groups vs. control. Column 3 presents the difference in means between the Positive and Balanced treatment groups. All 20 t-tests fail to reject equality, as do the joint F-tests of all characteristics. Like all reported results, these are weighted for national representativeness, although the unweighted groups are also balanced on all characteristics.

Appendix Table A2-1 presents correlations between baseline WTP and observable characteristics. Men, democrats, environmentalists, those who report having taken steps to conserve energy, and those with higher discount factors have higher demand for CFLs. (The discount factors are the δ parameter in a β, δ model of present bias, as calibrated from hypothetical intertemporal tradeoffs in the post-experiment survey.) These correlations conform to our intuition and build further confidence that the differences in WTP are meaningful. However, renters and more present-biased (lower β) consumers do not have lower WTP for CFLs conditional on other observables. This provides no support for the hypotheses that agency problems and present bias play a role in lightbulb decisions.

3.4 Empirical Strategy and Results

Figure 1 shows the baseline and endline demand curve for CFLs. The control endline demand curve sits almost directly on top of the baseline demand curve, implying that the control interventions had little effect on relative demand for CFLs vs. incandescents. The treatment endline curve is shifted outwards, reflecting an increase in demand for the CFL. At equal prices, which approximately reflects market conditions, 76 percent of the treatment group chooses the CFL, against 65 percent of control.

Figure 2 presents a histogram of the within-subject changes in WTP between baseline and endline. About 90 percent of control group consumers either have exactly the same WTP or change by $2 or less. In treatment, there is a mass to the right of the figure, with 36 percent of people increasing WTP by between $1 and $10.

Denote X_i as participant i's vector of characteristics from Table 1, and T_i as an indicator for whether the household is in either of the two treatment groups. We estimate the average treatment effects of the informational interventions on endline willingness-to-pay v_i^1 using OLS with robust standard errors:

$$v_i^1 = \tau T_i + \gamma X_i + \varepsilon_i \tag{1}$$

Table 2 presents the results. Column 1 presents the unconditional difference in means, excluding the Endline-Only group for comparability with other columns that include baseline WTP. Column 2 adds the control for baseline WTP v_i^0. Column 3 is the exact specification from Equation (1), including individual characteristics. The sample size decreases in column 3 because at least one X characteristic is missing for 15 consumers, but the effects do not change statistically. In column 3, the informational intervention caused consumers' WTP for the CFL to increase by an average of about $2.32.

One potential concern is that treatment group consumers might wish to be internally consistent in their choices between baseline and endline (Falk and Zimmermann 2012). This could cause endline choices to be biased towards the baseline, unlike an experimental design that did not require consumers to state baseline choices. This in turn would bias treatment effects toward zero. The Endline-Only treatment group was included to test this. Column 4 includes only the Endline-Only and control groups, excluding the treatment group that made baseline choices. The estimates should be compared against Column 1, which similarly does not control for baseline WTP. The point estimate is not statistically different, although it is lower by $0.46 per package.[8]

Top-coding and bottom-coding of WTP mechanically influence the treatment effect. Consumers with baseline WTP equal to the maximum cannot reveal a post-treatment increase in WTP, and any consumers with baseline WTP equal to the minimum could not reveal a decrease in WTP. Because the treatment tends to increase WTP, the former effect should dominate, and the average treatment effect should be understated. Column 5 excludes consumers with top-coded or bottom-coded baseline WTP of $v_i^0 = \$15$ or $v_i^0 = -\$15$. The estimated effect increases to $3.23.[9]

[8] Two other tests confirm why this result holds. First, average post-treatment WTP does not differ statistically between the Endline-Only group and rest of the treatment group. Second, the shape of demand does not differ between these two groups: Appendix Figure A2-1 shows that the Endline-Only demand curve sits very close to the demand curve for the rest of the treatment group, while control group demand is very different. Statistical tests show that the share of consumers with $v_i^1 > v^+$ does not differ statistically at any level of v^+.

[9] This increase is consistent with Figure 4, which we discuss later. The figure shows that consumers with top coded baseline relative WTP (plus the small group with WTP of $9) have close to zero conditional average treatment

Relatedly, the assumed mean censored value of \$15 caps the increase in WTP that any consumer can reveal. Since a larger share of endline WTP is top-coded in treatment relative to control (29 percent vs. 16 percent), increasing this assumed value should increase the treatment effect. In unreported regressions where we alternatively assume mean censored values of \$12 (\$20) instead of \$15, the ATE decreases to \$1.99 (increases to \$2.88).

3.4.1 Demand Effects

With any experiment other than a natural field experiment, one might be worried about demand effects: that participants change their actions to comply with, or perhaps defy, the perceived intent of the study. We could have designed a lengthier or otherwise more complex experiment to obfuscate our objective of estimating the effects of information disclosure, but this would have added to the cost. If demand effects are present, the likely direction would be to increase treatment group post-intervention WTP, i.e. make the treatment effect more positive. Aside from pointing out that the likely sign of the bias would only reinforce our qualitative conclusions, we also address demand effects in three ways.

First, we designed the experiment to include the Balanced treatment group, which disclosed both positive and negative information about CFLs. Consumers in this group should be less likely to believe that the experimenters were purely trying to persuade them to purchase the CFL. If demand effects play a large role, effects of the Positive treatment should be inflated relative to the Balanced treatment. Column 6 of Table 2 includes an indicator for the Positive treatment group, showing that the effects do not differ statistically, and the point estimates are fairly similar. Because the effects do not differ between the two groups, we combine all treated consumers in all other parts of the analysis.

Second, demand effects are less likely if participants cannot identify the intent of the study. The post-experiment survey asked consumers what they thought the intent of the study was. Multiple responses were allowed. Table 3 presents the share of each group that gave each response. The two treatment groups responded similarly, although the Balanced group was more likely to report that the intent of the study was to "understand what features of lightbulbs are most important to people" and less likely to report that the intent was to "test how well people are able to quantify energy costs." Relative to control, both treatment groups were more likely to respond that the intent of the study was to ""understand why people buy incandescents vs. CFLs," "test how well people are able to quantify energy costs," "test whether ability to quantify energy costs affects purchases," and "test whether consumer education affects purchases." The control group was more likely to respond that the intent was to "understand the effects of price changes," "measure whether people make consistent purchases in similar situations," and "test whether the number of bulbs in

effect, precisely because there is no way for them to increase their WTP in the multiple price list. This illustrates graphically why excluding these consumers increases the treatment effect.

a package affects purchasing patterns." The dispersion of beliefs within groups suggests that there is not one obvious way in which demand effects might act.

Third, if demand effects are present, they should differentially affect people who are more able to detect the intent of the study and are more willing to change their choices given the experimenter's intent. One existing measure of these issues is the Self-Monitoring Scale, a battery of personality questions developed by Snyder (1974). Snyder writes that the scale is designed to identify individuals who "tend to express what they think and feel, rather than mold and tailor their self-presentations and social behavior to fit the situation."

From the set of standard Self-Monitoring Scale statements, we took the most relevant six:

- It's important to me to fit in with the group I'm with.
- My behavior often depends on how I feel others wish me to behave.
- My powers of intuition are quite good when it comes to understanding others' emotions and motives.
- My behavior is usually an expression of my true inner feelings, attitudes, and beliefs.
- Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.
- I would NOT change my opinions (or the way I do things) in order to please someone else or win their favor.

At the very end of the post-experiment survey, we asked consumers to respond to each of these six statements on a five-point Likert scale from "Agree" to "Disagree." We normalize responses to each question to mean zero, standard deviation one, and interact each with the treatment indicator while also controlling for lower-order interactions. While the six Self-Monitoring Scale variables are correlated with each other, none is correlated with endline CFL demand or with the treatment effect, nor is a composite of the six. The estimation results can be found in Appendix Table A2-2.

3.4.2 Mechanisms

How much of the treatment effect is coming from changing information sets vs. directing attention to energy costs? The post-experiment survey elicits beliefs over how much less it costs to buy electricity for a CFL vs. incandescents over the typical 8000-hour life of a CFL, at national average electricity prices. The question is similar, but not identical, to the "quiz" question asked of the treatment group, and the correct answer is \$36 (\$48 for the incandescent minus \$12 for the CFL). Column 1 of Table 4 shows that the treatment increases median beliefs: they are \$25 in control and \$13 higher in treatment. Column 2 of Table 4 shows that the treatment also substantially reduces the median absolute error, i.e. the absolute value of the difference between the reported belief and \$36. As in the other tables, the exact sample sizes are slightly smaller than the total number of qualified participants because a handful of participants refused to answer. We use median regressions because reported beliefs have high variance, and median regressions are more robust to extreme outliers than OLS. These results suggest that at least part of the treatment effect is from changing information sets.

The post-experiment survey also asks consumers to rate on a scale of 1-10 the importance of energy use, bulb lifetime, warm-up time, and mercury and disposal in their purchase decisions. Table 5 presents how the treatments affected these ratings. Both Positive and Balanced treatments decreased the stated importance of purchase prices, consistent with consumers re-orienting away from purchase price as a measure of cost. Point estimates suggest that both the Positive and Balanced treatments increased the importance of energy use and that the Positive treatment also increased the importance of bulb lifetimes. These are the only estimates in the entire analysis whose significance level is affected by the weighting: they are not significant in Table 5, but (unreported) regressions show that they are statistically significant when weighting all observations equally. The Positive treatment and control groups do not differ on the importance of warm-up time or mercury and disposal, which is to be expected because neither group received information on these two issues. Interestingly, the Balanced treatment decreased the importance of warm-up time. One potential explanation is that consumers had previously believed that CFL warm-up times were longer, and the treatment reduced the importance of this difference between CFLs and incandescents.

4 In-Store Experiment

4.1 Experimental Design

Would the effects of information provision be different in a more typical retail setting compared to the TESS platform? To answer this, we partnered with a large home improvement retailer to implement an in-store experiment. Between July and November 2011, three research assistants (RAs) worked in four large "big box" 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.[10] They told customers that they were from Harvard University and asked, "Are you interested in answering some quick research questions in exchange for a discount on any lighting you buy today?" 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, the number of bulbs they were buying, and the amount of time each day they expected these lightbulbs to be turned on each day. The survey did not mention electricity costs or discuss any differences between incandescents and CFLs.

The iPad randomized customers into treatment and control groups with equal probability. For the treatment group, the iPad would display the annual energy costs for the bulbs the customer

[10] This includes standard bulbs used for lamps and overhead room lights. Specialty bulbs like Christmas lights and other decorative bulbs, outdoor floodlights, and lights for vanity mirrors are sold in an adjacent aisle.

was buying, given his or her estimated daily usage. It also displayed the total energy cost difference over the bulb lifetime and the total user cost, which included energy costs plus purchase prices. Appendix 3 presents the information treatment screen. The RAs would interpret and discuss the information with the customer, but they were instructed not to advocate for a particular type of bulb and to avoid discussing any other issues unrelated to energy costs, such as mercury content or environmental externalities. The control group did not receive this informational intervention, and the RAs did not discuss energy costs or compare CFLs and incandescents with these customers.

At the end of the survey and potential informational intervention, the RAs gave customers a coupon in appreciation for their time. The iPad randomized respondents into either the Standard Coupon group, which received a coupon for 10 percent off all lightbulbs purchased, or the Rebate Coupon group, which received the same 10 percent coupon plus a second coupon valid for 30 percent off all CFLs purchased. Thus, the Rebate Coupon group had an additional 20 percent discount on all CFLs. For a consumer buying a typical package of 60 Watt bulbs at a cost of $3.16 per bulb, this maps to an average rebate of $0.63 per bulb. The coupons had bar codes which were recorded in the retailer's transaction data as the customers submitted them at the register, allowing us to match the iPad data to purchases.

After giving customers their coupons, the RAs would leave the immediate area in order to avoid any potential external pressure on customers' decisions. The RAs would then record additional demographic information on the customer, including approximate age, gender, and ethnicity. The RAs also recorded this information for people who refused. Finally, the RA recorded the total duration of the interaction. The difference between treatment and control had a mean of 3.17 minutes and a median of 3.0 minutes. This measures the amount of time spent discussing the energy cost information and the differences between incandescents and CFLs.

4.2 Data

Of the 1561 people who were approached, 459 refused, while 1102 began the iPad survey. Of these, 13 broke off after the first question, two broke off later, and 1087 were assigned to a treatment group and given a coupon. Column 1 of Table 6 presents descriptive statistics for the sample of customers who completed the survey and were given a coupon. Column 2 presents differences between the 474 people who refused or did not complete the survey and the 1087 who completed, using the demographic characteristics recorded for those who refused. People whom the RAs thought were older, male, Asian, and Hispanic were more likely to refuse. Columns 3 and 4 present differences between the information treatment and control groups and between the rebate and standard coupon groups. In one of the 18 t-tests, a characteristic is statistically different with 95 percent confidence: we have slightly fewer people coded as Asian in the information treatment group. F-tests fail to reject that the groups are balanced.

We restrict our regression sample to the set of consumers that purchase a "substitutable light-bulb," by which we mean either a CFL or any incandescent or halogen that can be replaced with a CFL. The bottom panel of Table 6 shows that 77 percent of interview respondents purchased any lightbulb with a coupon, and 73 percent of survey respondents purchased a substitutable lightbulb. While information or rebates theoretically could affect whether or not customers purchase a substitutable lightbulb, t-tests show that in practice the percentages are not significantly different between the groups.

4.3 Empirical Strategy and Results

We denote T_i and S_i as indicator variables for whether customer i is in the information treatment and rebate groups, respectively. X_i is the vector of individual-level covariates. We estimate a linear probability model[11] with robust standard errors using the following equation:

$$1(\text{Purchase CFL})_i = \eta S_i + \tau T_i + \alpha X_i + \varepsilon_i \tag{2}$$

Table 7 presents estimates of Equation (2). Column 1 excludes covariates X_i, while column 2 adds them. The estimates are statistically identical, and the point estimates are very similar. For customers who received the standard coupon and were in the information control group, the CFL market share is 34 percent. The rebate increases CFL market share by about ten percentage points. This implies a price elasticity of demand for CFLs of $\frac{\Delta Q/Q}{\Delta P/P} \approx \frac{0.1/0.34}{-0.2} \approx -1.5$. Column 3 shows that the interaction between information and rebates is statistically zero.

The informational intervention does not statistically affect CFL market share. Using the standard errors from column 2, we can reject with 90 percent confidence that the intervention had more than 73 percent of the effect of the 20 percent CFL rebate. Assuming linear demand, this bounds the effect of information at the effect of a 12 percent rebate, or about \$0.46 per 60-Watt equivalent bulb.

There are several reasons for why the information effect might differ from the TESS experiment. First, it could be that a very large share of consumers did not understand the in-store informational intervention or were in too much of a hurry to internalize the information. However, our RAs reported that they believe that this is unlikely. Second, the informational environment differs: these and other home improvement stores have signage in lightbulb aisles that highlights features of different lightbulb technologies, including electricity use. If this existing information is very effective, incremental information might have no effect. Notice that if this is the case, our treatment effects are still the relevant parameters for policy analysis later in the paper: if existing information provision technologies are fully effective, then there is no remaining imperfect information and inattention to justify subsidies and standards.

A third reason is that the experimental populations differ: the TESS population is nationwide, while the in-store sample is drawn from four stores in three eastern states. Home improvement

[11] We technically prefer the linear probability model here because we assume locally linear demand when using the estimates for policy analysis. In any event, S and T are indicator variables, and the probit estimates are almost identical.

retailers are the most common place where households buy lightbulbs (DOE 2010), and our partner alone sells upwards of 50 million lightbulb packages each year, a non-trivial share of national sales. Internally valid estimates for our experimental sample are thus of great interest *per se*, even if the results might not generalize to other types of retailers.

5 A Framework for Policy Analysis

5.1 Consumers

We model consumers that make one of two choices, labeled E and I. In our empirical application, E represents the purchase of an energy efficient product (the CFL), while I is an energy inefficient product (the incandescent). More generally, this model could capture any choice over which consumers might misoptimize.

Products $j \in \{E, I\}$ are sold at prices p_j, and $p = p_E - p_I$ is the relative price of E. We define v_j as the consumer's true utility from consuming product j, and we call $v = v_E - v_I$ the relative true utility from E. In our empirical application, v could be determined by any and all of the differences between CFLs and incandescents, such as electricity costs, longer lifetime, mercury content, brightness, and "warm glow" utility from reduced environmental impact.

A consumer's utility from purchasing product j at price p_j is given by $v_j + (Z - p_j)$, where Z is the consumer's budget and $Z - p_j$ is utility from consumption of the numeraire good. A fully optimizing consumer thus chooses E if and only if $v > p$. A misoptimizing consumer chooses E if and only if $v - b_k(p) > p$, where $b_k(p)$ is a bias that may depend on p and is continuously differentiable in p. To simplify notation, we will typically denote bias by b_k rather than $b_k(p)$. The pair (v, b_k), which we will refer to as a consumer's "type," is jointly distributed according to a distribution F. We denote the conditional distribution of v given b by $F_{v|b_k}(\cdot|b_k)$. We assume that for each b_k, the distribution $F_{v|b_k}(\cdot|b_k)$ has an atomless and continuous density function $f_{v|b_k}(\cdot|b_k)$. For simplicity, we will also assume that b takes on finitely many values, though the analysis easily generalizes. We let α_k denote the fraction of consumers with bias b_k.

We also call b_k the "internality," to highlight the analogy to externalities. While an externality is a wedge between private willingness-to-pay (WTP) and social welfare, the internality is a wedge between private WTP and true private welfare. This is a reduced form representation of many biases that could cause consumers not to maximize experienced utility, including misperceptions of any product attribute. It allows for dependencies between bias b_k, true valuation v, and price p, as theories of endogenous inattention would imply. In our empirical application, we think of the bias as arising from consumers' undervaluation of energy costs due to a set of informational and attentional biases that we discuss in the next section.

Under the additional assumption below, this model generates continuous and downward-sloping demand curves for product E.[12]

Assumption 1: b_k is differentiable in p and there exists a $\rho > -1$ such that $b_k'(p) > \rho$ for all p.

Let $D^R(p) = 1 - F(p)$ denote the "unbiased" demand curve for E, let $D_{b_k}(p) = \alpha_k[1 - F_{v|b_k}(p + b_k|b_k)]$ denote the demand curve of consumers with bias b_k, let $D^R_{b_k} = \alpha_k[1 - F_{v|b_k}(p|b_k)]$ denote what would be the demand curve of consumers with bias b if they were "debiased," and let $D(p) = \sum_k D_{b_k}(p)$ denote the total demand curve of all consumers. Our assumptions about b_k and F imply that all demand curves are continuously differentiable functions of p. All analysis that follows expresses results in terms of these demand curves, so this framework could also be applied to continuous choice situations.

5.2 The Policymaker

The policymaker has two types of tax policies available: a subsidy of amount s for E and a ban on either choice. We will compare the welfare impacts of these policies to a hypothetical technology that can fully debias consumers.

The policymaker maintains a balanced budget through lump-sum tax or transfer $T(s) = \int s\sigma(v, b_k, c - s)dF(v, b_k)$. This implies that the subsidy has no distortionary effects on other dimensions of consumption, and thus its role is purely corrective. Because all consumers choose either E or I, the subsidy for E is equivalent to a tax on I, and a ban on one choice is equivalent to a mandate for the other. Products E and I are produced in a competitive economy at a constant marginal costs c_j, with relative cost $c = c_E - c_I$. Product E's relative price after subsidy s is $p = c - s$.

Let $\sigma(v, b_k, p)$ denote the choice choice of a type (v, b_k) consumer at price p, with $\sigma = 1$ denoting the choice of E and $\sigma = 0$ denoting the choice of I. Also denote p_σ as p_E if $\sigma = 1$ and p_I if $\sigma = 0$. Finally, define a normalizing constant C as the integral over all consumers of $Z + v_I - p_I$. The policymaker's objective is

$$\begin{aligned} W(s) &\equiv \int [\overbrace{\sigma(v, b_k, c - s)(v)}^{\substack{\text{relative true utility} \\ \text{from } E}} \overbrace{+Z + T(s) - p_\sigma}^{\substack{\text{consumption of} \\ \text{numeraire good}}}]dF(v, b_k) \\ &= C + \int \sigma(v, b_k, c - s)(v - c)dF(v, b_k). \end{aligned} \tag{3}$$

Our assumptions about b_k and F imply that W is continuously differentiable. To ensure the existence of an optimal subsidy, we assume that set of possible values of $v - b_k$ is bounded from above and from below; that is, $\bigcup_k \mathrm{supp} F_{v|b_k}(\cdot|b_k)$ is a bounded set. This assumption implies that a

[12] The assumption that $b_k'(p) > -1$ is needed to gurantee that a consumer's perceived relative value of E, given by $v - b_k(p) - p$, does not increase in p. The additional assumption that b_k' is bounded away from -1 guarantees that $v - b_k(p) - p < 0$ for high enough p and that $v - b_k(p) - p > 0$ for low enough p, which implies that type k is marginal at some price.

ban on product I is equivalent to a sufficiently large subsidy, and a ban on product E is equivalent to a sufficiently small (negative) subsidy (i.e., sufficiently large tax on E).

An important benchmark we will consider is the *first best level of welfare*, given by

$$W^{FB} = C + \int_{v \geq c} (v - c) dF. \tag{4}$$

5.2.1 Subsidy

Define the *average marginal internality* at price $p = c - s$ as

$$B(p) = \frac{\sum_k b_k D'_{b_k}}{D'}.$$

In Appendix 4, we establish the following result, which is analogous to the optimal tax formulas derived by Allcott, Mullainathan and Taubinsky (2013) in a similar setting:

Proposition 1

$$W'(s) = (B(c - s) - s) D'(c - s) \tag{5}$$

The intuition behind Proposition 1 is that the welfare impact of a subsidy trades off the internality reduction, $B(c - s)D'$, with the distortion to consumers' decision utility, $-sD'$. This is directly analogous to the logic behind a Pigouvian tax, which trades off externality reduction with distortions to consumers' private utility gains. Allcott, Mullainathan and Taubinsky (2013) discuss this parallel between internalities and externalities in more depth.

Since an optimal subsidy s^* must satisfy $W'(s^*) = 0$, Proposition 1 implies that an optimal subsidy must equal the average marginal internality:

$$s^* = B(c - s^*) \tag{6}$$

Equation (6) is analogous to Diamond's (1973) result that the optimal externality tax when agents have heterogeneous externalities is a similarly-weighted average marginal externality.

Figure 3 illustrates the intuition in the special case with linear demand and constant $B(p)$. The unbiased demand curve $D^R(p)$ is shifted out relative to demand curve $D(p)$, meaning that the bias reduces demand for good E and causes welfare losses. The average marginal internality is the distance from a to f. The initial equilibrium is at point b, and the optimal subsidy s^* moves the equilibrium to point f. A marginal increase in the subsidy from 0 induces marginal consumers at point b to purchase good E, increasing their true utility by amount bd. The welfare gain from the optimal subsidy is triangle abd.

Equation (6) highlights the kinds of consumer heterogeneity that do and do not matter for the optimal subsidy. Heterogeneity in the average marginal internality $B(p)$ at different price levels clearly does matter, and there are a number of practical situations where one might expect higher-WTP consumers to be more or less biased. For example, environmentalist consumers may both be more attentive to energy costs and have higher true relative utility v from good E. This highlights that it is not sufficient to set an internality tax based on a general idea that "consumers are biased" - it matters whether the biased consumers are marginal to the policy. As the analogy to Diamond's (1973) formula suggests, this insight generalizes to other market failures: if setting a time-invariant congestion tax, for example, the optimal tax would be smaller if people who travel at rush hour (and thus impose larger externalities) are less price elastic.

However, heterogeneity in bias b *within* the set of consumers on the margin at price p does not affect the optimal subsidy: only the average marginal internality matters. This is important because some models such as Chetty, Looney, and Kroft (2007) have consumers that are either fully unbiased or completely biased with some probability, while other models might have all consumers with a partial bias. While empirical analyses such as Chetty, Looney, and Kroft (2009), Hossein and Morgan (2006), and Abaluck and Gruber (2011) have been able to identify average biases within groups of marginal consumers, we are not aware of previous studies that have been able to identify distributions of individual biases. Equation (6) shows that the optimal subsidy can be set without knowledge of the underlying "structural" model and distribution of biases. Thus, the $B(p)$ function is a sufficient statistic for setting the optimal subsidy in the sense of Chetty (2009).

While the $B(p)$ function is a sufficient statistic for calculating the welfare impacts of a subsidy, it is not informative about how close the optimal subsidy comes the first best. As the next proposition shows, a subsidy can achieve the first best if and only if consumers have homogeneous bias. Combined with the fact that the same $B(p)$ function can be generated by either homogeneously or heterogeneously biased consumers, this proposition implies that $B(p)$ is not informative of the gap between the second best welfare attained by the optimal subsidy and the first best level of welfare that would be attained if consumers were fully debiased.

Proposition 2 *Let s^* be an optimal subsidy. If all consumers have the same bias b, then s^* is uniquely defined and $W(s^*) = W^{FB}$. If some consumers have bias b_i while other consumers have bias b_j such that $b_i(p) < b_j(p)$ for all p, then $W(s^*) < W^{FB}$.*

A key implication of Proposition 2 is that if an informational intervention that fully debiases consumers were inexpensive and feasible at large scale, it would likely be preferred to a subsidy or ban. If bias b is heterogeneous, a subsidy is less efficient: if it perfectly corrects the choice of consumers with bias b_1, it still leaves consumers with bias $b_2 > b_1$ to underpurchase E, while leaving consumers with bias $b_3 < b_1$ to overpurchase E. This is the intuition for why "asymmetric paternalism" (Camerer *et al.* 2003) and "libertarian paternalism" (Sunstein and Thaler 2003) are preferred to subsidies and bans. The reason to also consider subsidies and bans is if fully-debiasing information provision technologies are costly or infeasible, while the feasible information provision technologies do not fully debias consumers. In particular, our informational interventions are unlikely to be scaled: even though we chose high-volume store locations, the in-store experiment required labor costs of several dollars per customer intercept, and it is not obvious how the TESS intervention could be implemented outside the TESS platform.

5.2.2 Ban

According to welfare equation (3), a ban on good I has the following effect on welfare:

$$\begin{aligned}\Delta W &= \int (v-c)dF(v,b_k) - \int \sigma(v,b_k,c)(v-c)dF(v,b_k) \\ &= \int (1-\sigma(v,b_k,c))(v-c)dF(v,b_k) \\ &= \int_{\{(v,b_k)|\sigma(v,b_k,c)=0\}} vdF(v,b_k) - c \end{aligned} \tag{7}$$

This equation simply states that the welfare effects of a ban are the average relative true utility v for consumers currently purchasing I, minus the relative cost c. An analogous equation would hold for a ban on E.

Figure 3 illustrates this equation, again assuming linear demand and constant $B(p)$. The ban on good I increases welfare for the set of consumers to the left of point f, because purchasing good E increases their true utility. This welfare gain is the triangle abd. However, the ban decreases welfare for the set of consumers to the right of point f: while they are biased, their true utility from good E is still less than the relative price. This welfare loss is the triangle amn.

Under our assumption of lump-sum revenue recycling, a ban on good I is equivalent to a sufficiently large subsidy for good E. Bans are thus weakly worse than the optimal subsidy. However, there is some marginal cost of public funds at which a ban could be preferred. On the other hand, if the corrective price policy is implemented as a tax on good I, then a cost of public funds further reinforces the relative appeal of price-based policies relative to bans.

5.3 First-Order Approximation to Optimal Subsidy

When the policymaker can directly measure the $B(p)$ function at all p, welfare changes can be computed at each subsidy level to exactly compute the globally optimal subsidy. In Section 6, we use the TESS experiment to do this. However, this function is in general difficult to estimate: there are only a few papers in any context that cleanly identify biases for even some subset of consumers. We now present one approach to approximating the marginal internality using two reduced form sufficient statistics: the slope of demand and the effect of the bias on market shares. In Section 6, we illustrate how this can be implemented using the in-store experiment.

To a first order approximation, we have

$$D_{b_k}^R(p) - D_b(p) = D_{b_k}(p - b_k) - D_{b_k}(p)$$
$$\approx b_k D'_{b_k}(p).$$

Thus,

$$D^R(p) - D(p) = \sum_k (D_{b_k}(p - b_k) - D_{b_k}(p)) \approx \sum_k b_k D'_{b_k} = D'(p)B(p).$$

It then follows that

$$B(p) \approx \frac{D^R(p) - D(p)}{D'(p)}. \tag{8}$$

The numerator is the effect of the bias on market shares, while the denominator is the slope of demand. In other words, the average marginal internality is the price change that affects quantity demanded exactly as much as the bias does.

To a first-order approximation, demand and "unbiased demand" have the same slope: $(D_{b_k}^R)'(p) = D'_{b_k}(p - b_k) \approx D'_{b_k}(p)$. This means that we can also approximate B by

$$B(p) \approx \frac{D^R(p) - D(p)}{(D^R)'(p)}. \tag{9}$$

Figure 3 illustrates the intuition behind Equation (8). The length of segment ab is given by $D^R(p) - D(p)$. This could be identified through a randomized field experiment that fully debiases consumers and measures the effects on demand for E. The demand slope $D'(p)$ could be identified through an RCT that randomizes relative prices. The segment af, which corresponds to the average marginal internality, is given by $\frac{D^R(p)-D(p)}{D'(p)}$. Combining Equation (8) or (9) with Equation (6) yields an approximation to the optimal subsidy. Thus, the effect of the bias on market shares and the slope of demand are sufficient statistics for an approximation to the average marginal internality.

6 Policy Evaluation

6.1 Inferring Bias from Treatment Effects

We now combine the experimental estimates with the theoretical framework to illustrate how results such as these could be used to inform policy. To do this, we build on the idea that the information treatment groups choose optimally, although the information control group may not. One important feature of this approach is that it "respects choice" in the sense of Bernheim and Rangel (2009): we conduct welfare analysis using consumers' own choices in what is plausibly a "debiased" state. In their language, we define control group choices as *provisionally suspect* due to the possibility of imperfect information processing. If choices differ between treatment and control, we delete control group choices from the welfare-relevant domain. In our language, the implication is that the conditional average treatment effect of our informational interventions at any price p equals the average marginal internality from imperfect information and inattention:

Assumption 2: $\tau(p) = B(p)$.

This is analogous to the assumption made by Chetty, Looney, and Kroft (2009) when they estimate the magnitude of inattention to sales taxes using the treatment effect of an intervention that posted tax-inclusive purchase prices. In justifying this assumption, Chetty, Looney, and Kroft (2009) write that "when tax-inclusive prices are posted, consumers presumably optimize relative to the tax-inclusive price." Similarly, it seems reasonable to assume that consumers optimize relative to lightbulb lifetimes and energy costs after we provided them with information about these attributes. In the empirical sections, however, we have discussed potential reasons why this assumption may not hold. In qualitatively interpreting the results, we view this assumption as an approximation.

6.2 "Structural" Models of Bias

Because the optimal policy depends on b, not the underlying "structural" model of the bias, our exposition uses this "reduced form" parameter. However, it may be helpful to specify categories of inefficiencies that could affect lightbulb demand and would plausibly be addressed by our informational interventions:

1. Costly information acquisition, as in Gabaix *et al.* (2006) and Sallee (2013). This category includes many standard models of imperfect information in which the consumer incurs a cost to learn about energy efficiency, lifetime, or other product attributes and, in the absence of paying that cost, assumes that different goods have the same level of an attribute.

2. Biased priors about energy costs or other product attributes, as tested by Allcott (2013), Attari *et al.* (2010), Bollinger, Leslie, and Sorensen (2011), and others. Put simply, this category reflects consumers who may know that CFLs use less energy but don't know that the savings are so large.

3. Exogenous inattention to energy as a "shrouded" add-on cost, as in Gabaix and Laibson (2006).

4. Costly cognition or "thinking cost" models, as in Conlisk (1988), Chetty, Looney, and Kroft (2007), Gabaix (2013), Reis (2006), Sims (2004), and others. In these models, consumers might not pay attention to differences in energy costs between lightbulbs because their experiences with other goods suggest that energy cost differences are typically unimportant. However, once informed that lightbulb energy cost differences are large relative to the difference in purchase prices, consumers in these models would consider them in their choices.

Informational interventions would not affect all biases that could affect lightbulb demand. For example, "bias toward concentration" (Koszegi and Szeidl 2013) could cause consumers to undervalue electricity costs because they occur in a stream of small future payments. Koszegi and Szeidl

(2013) point out that re-framing the stream of payments as one net present value, as our interventions do, does not necessarily address this possible bias. It is also possible that present bias over cash flows could cause consumers to undervalue the CFL's future cost savings, although this is not consistent with the TESS data or the standard models of present bias over consumption. Our informational interventions should not affect present biased consumers. Finally, consumers could be imperfectly informed about or inattentive to other attributes not discussed in our informational interventions.

Because imperfect information and inattention may not be the only biases, Appendix 4 generalizes the theoretical framework to the case when the intervention identifies only part of the bias. Intuitively, the optimal subsidy is additive in the different types of internality. For example, if present bias over cash flows reduces marginal consumers' demand for the CFL, then the true optimal subsidy is larger than the subsidies we calculate based on the informational interventions alone. Similarly, if present bias reduces inframarginal consumers' CFL demand, then the true welfare gains of a ban are larger (or less negative) than we calculate.

6.3 Using the TESS Experiment Results

6.3.1 Subsidy

The theoretical framework shows that to set and evaluate policy, we need to know both the initial demand curve and the average marginal internality at each point. This should now clarify why the particular design we used for the TESS experiment is so important: by eliciting WTP in consumers' baseline (potentially biased) state and subsequently in their treated (optimizing) state, we can identify the average marginal internality at each point on the market demand curve.

Figure 4 presents the conditional average treatment effects (CATEs) at each level of baseline WTP. As Figure 1 shows, there are only a small number of consumers with baseline WTP equal to $9 or between -$3.50 and -$9, so we group outlying high and low baseline WTPs together. Consistent with the ATEs in Table 2, the CATEs are all around $3, except for the CATE at the highest baseline WTP, which is close to zero. This is simply due to top-coding: consumers who start with top-coded baseline WTP cannot increase their WTP further. Because these inframarginal consumers are unaffected by the subsidy and the ban, this does not affect the welfare calculations. After excluding consumers with top-coded and bottom-coded baseline WTP, there is a slight positive correlation between baseline WTP and the treatment effect. This highlights that the population average internality would not be the right parameter for setting optimal policy.

Figure 5 illustrates how this distribution of average marginal internalities is combined with the baseline demand curve for policy analysis. The dashed line is the baseline demand curve $D(p)$. At each point, the average marginal internality from Figure 4 is added to WTP to give the average true utility of consumers marginal at each price p. These average true utilities are plotted as diamonds. Consistent with their approximately equal retail prices, we assume that the two packages have the same marginal cost of production, so $c = 0$.

The leftmost shaded rectangle reflects the welfare gain from increasing the subsidy from \$0 to \$1. This increased subsidy moves about 13 percent of consumers over the margin to buying a CFL. The height of the rectangle is the difference between average true utility and relative cost c. Moving to the right, the next two shaded rectangles reflect the welfare gain from increasing the subsidy from \$1 to \$2 and from \$2 to \$3, respectively. These first three shaded rectangles reflect welfare gains, as average true utility v exceeds cost c. However, further increases in the subsidy cause welfare losses. The \$3 optimal subsidy is consistent with Figure 4, which shows ATEs in the range of \$2 to \$3 for consumers with baseline WTP less than 0.

The welfare effects of banning incandescents are the sum of all shaded rectangles. Notice again that it is the average marginal internalities that determine welfare impacts, not the distribution of individual biases within the sets of marginal consumers. Graphically, this is reflected by the fact that height of each discrete welfare rectangle is determined only by the average internality at each price.

Table 8 presents formal calculations that parallel Figure 5. Column 1 contains the subsidy amount. Column 2 presents the average baseline relative WTP v^0 for consumers marginal to the increase in the subsidy, assuming that demand is linear between the two price levels. Column 3 presents the average internality for this group of marginal consumers. This equals the treatment effect for each point on the left side of Figure 4. Column 4 presents the demand density: the share of all consumers that are marginal at each relative price level. Column 5 presents the welfare effect of the increment to the subsidy, using Equation (5), while column 6 presents the total welfare effect of changing the subsidy from zero to the amount listed in that row. Columns 3-6 are measured with sampling error, although we omit standard errors for simplicity.

As Equation (5) shows, a marginal increase in the subsidy increases welfare as long as the marginal internality outweighs the distortion to decision utility. Table 8 shows that for subsidies larger than \$3, the point estimate of true utility for marginal consumers is less than the cost of the CFL. Thus, increases in the subsidy above \$3 will reduce welfare relative to the \$3 subsidy in our model. Using the analogy to externalities, subsidies higher than \$3 would be equivalent to setting a Pigouvian externality tax higher than marginal damages. For comparison, typical CFL rebates offered by electric utilities have been on the order of \$1 to \$2 per bulb.

If we assume that $B(p) = \widehat{\tau}(p) = \widehat{\tau} = \$2.32$ at all p, Equation (6) implies that the globally optimal subsidy is also $s^* = \$2.32$, which is consistent with the result in Table 8 that the optimal subsidy does not exceed \$3. However, the benefit of this "grid search" approach to calculating the optimal subsidy is that in theory, the average marginal internality could be very different for different values of v^0. Grid search identifies the global optimum even if the necessary condition for a local optimum in Equation (6) is satisfied at multiple subsidy levels.

6.3.2 Ban

A ban on traditional incandescents is equivalent to a change in relative prices that is so large as to induce all consumers to cease buying incandescents. In Figure 5, this is the sum of the positive welfare rectangles above the x-axis minus the sum of the negative welfare rectangles below the x-axis. Table 8 shows that this sums to a loss of \$0.436 per package sold.

Top-coding and bottom-coding have two opposing effects on this welfare calculation. First, the treatment causes many treatment group consumers to be willing to pay the maximum for the CFL. Assuming a larger average WTP for this top-coded group would increase the treatment effect, implying a larger bias and thus larger welfare gains from corrective policies. Second, however, the welfare effects of a ban depend importantly on the tail of the WTP distribution: if some consumers very strongly prefer incandescents, banning them can cause large welfare losses. Appendix Table A2-3A tests the sensitivity of these welfare calculations to assuming that top-coded and bottom-coded relative WTPs average \$12 and -\$12. The two opposing effects almost exactly offset each other: the welfare loss from the ban is \$0.434 per package sold. Appendix Table A2-3B instead assumes mean censored WTPs of \$20 and -\$20. Under this assumption, the welfare loss from the ban is \$0.744 per package sold.

6.3.3 Illustrative Calculation: Welfare Gains from Information Provision

Proposition 2 shows that when the bias is heterogeneous, a fully-debiasing informational intervention generates larger consumer welfare gains than a subsidy. For the purpose of illustrating this, we briefly make a very strong assumption: that each treated consumer's individual WTP change from baseline to endline equals b. This is stronger than our Assumption 2, which was that the conditional average treatment effects equal $B(p)$.

Debiasing changes a consumer's welfare if and only if it changes his or her choice at market prices. The utility gain for a consumer who does change his or her choice is $|v-c|$. Thus, the welfare gain from full debiasing is the integral of $|v-c|$ over all consumers who change choices:

$$\int |\sigma(v,0,c)-\sigma(v,b_k,c)|\cdot|v-c|dF(v,b_k). \tag{10}$$

In the TESS data, 21 percent of treatment group consumers change choices after the intervention, and their average $|v-c|$ is \$3.36, giving total welfare gains of \$0.72 per package.[13] This is almost three times larger than the \$0.26 per package welfare gain from the optimal subsidy. Of course, to fully compare the two policies, one would need to subtract the cost of implementing each

[13] We emphasize that this calculation is purely for the purposes of illustrating Proposition 2. Figure 2 illustrates why the required assumption is too strong: some control group consumers also change WTP between baseline and endline, even though the control intervention was not designed to debias. Under a similar assumption, Equation (10) would imply that the welfare gains from the control intervention are \$0.19. This suggests that even if the conditional average treatment effects are meaningful, there can be noise in any given consumer's WTP change between baseline and endline.

policy, including the costs of consumers' time for the informational intervention and any deadweight loss of raising public funds for the subsidy. This calculation simply illustrates the sense in which the "targeting" properties of information provision can make it preferred to a subsidy.

6.4 Using the In-Store Experiment Results

The treatment effects from the in-store experiment can be used in Equation (8) to determine the optimal subsidy for this sample in this context. If the treatment group optimizes with respect to energy costs and product lifetimes, while the control group is potentially biased, $\widehat{\tau} = D^R(p) - D(p)$. We assume that $D' = (D^R)'$, as we do not reject this hypothesis in Table 7, and it is theoretically true to first-order approximation. Plugging $\widehat{\tau}$ and $\widehat{\eta}$ from column 2 of Table 7 into Equation (8) and using that the average rebate s per 60-Watt bulb was \$0.63, the optimal subsidy per 60-Watt bulb is:

$$B(S=0) \approx \frac{D^R(p) - D(p)}{D'(p)}. = \frac{\widehat{\tau}}{\widehat{\eta}/s} \approx \frac{0.004}{0.105/\$0.63} \approx \$0.024 \tag{11}$$

Figure 6 illustrates the calculation. The information treatment effect $\widehat{\tau}$ is the distance from b to a, and the slope of demand is $\widehat{\eta}/s$. The figure exaggerates $\widehat{\tau}$, as the point estimate suggests only a very small effect on market share.

Using the Delta method and the estimated variance-covariance matrix, the 90 percent confidence interval on the optimal subsidy per 60-Watt equivalent CFL is $(-\$0.30, \$0.35)$. Thus, for this sample of people in the informational environment where our experiment took place, the results rule out that the optimal subsidy is more than one-third as large as the \$1 to \$2 per bulb subsidies that electric utilities have typically offered. Furthermore, unless the treatment group consumers who bought incandescents in this experiment have substantially weaker preferences for incandescents than the inframarginal consumers in the TESS experiment, the tighly-estimated zero treatment effect suggests that banning incandescent lightbulbs will cause larger welfare losses in this population than in the TESS population. Thus, while the empirical estimates from the two experiments are different, they both lead to the same qualitative conclusion about the ban.

7 Conclusion

Imperfect information and inattention are commonly-proposed justifications for energy efficiency subsidies and standards, and they are frequently invoked in the lightbulb market as justifications for energy efficiency policies. We implemented two randomized control trials that measure the effects of "powerful" information provision on purchases of energy efficient CFLs. The TESS intervention increased WTP by an average of \$2.32, while the in-store experiment had tightly-estimated zero effects in a different population and informational environment. These forms of information provision would be difficult to scale, and lower-cost disclosure technologies seem less likely to fully inform and debias consumers. We thus formalized a theoretical framework that uses the experimental results to evaluate two second best policies: CFL subsidies and a ban on traditional incandescent bulbs. Results suggest that moderate CFL subsidies may be optimal, but that imperfect information and inattention do not justify a ban on traditional incandescents.

For our quantitative evaluation of subsidies and standards, we assumed that the treatment effects of powerful information provision identify the magnitude of bias. Given our experimental designs, we think that the assumption is a reasonable but imperfect approximation. We have discussed the importance of issues such as whether the treatment groups understood the information, experimenter demand effects, and external validity.

Even though the policy analysis is approximate, there are several ways in which this study is valuable. First, the in-store experiment is a proof-of-concept for what we think will be an important research effort to use field experiments to evaluate the effects of energy efficiency subsidies and information provision on purchases of durable goods. Second, we have highlighted the necessary parameters for studying the "internality rationale" for energy efficiency policies, and we have implemented two examples of experimental designs that can identify these parameters. Third, as we calculated in Section 2, estimates of inattention from other research combined with the very large magnitude of lightbulb energy costs relative to purchase prices suggested that biases from imperfect information and inattention might have had large effects in this market. Results from both experiments are consistent in that they reject that potential prior. Fourth, our model provides one initial data point which suggests that imperfect information and inattention may not justify the lighting energy efficiency standards in the absence of other distortions.

Although our application is to one particularly important and controversial policy, the approach is quite general. The theoretical framework generalizes immediately to any binary or continuous choice, and the idea of using informational interventions in RCTs to quantify internalities can be used in a wide variety of contexts. The approach to optimal policy could be useful in other contexts where informational or attentional biases might be used to justify subsidies or bans and where powerful information provision is feasible in small RCTs but not cost effective at large scale.

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

[3] Allcott, Hunt (2013). "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy*, Vol. 5, No. 3 (August), pages 30-66.

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

[5] Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky (2013). "Energy Policy with Externalities and Internalities." *Journal of Public Economics*, forthcoming

[6] Allcott, Hunt, and Nathan Wozny (2013). "Gasoline Prices, Fuel Economy, and the Energy Paradox." *Review of Economics and Statistics*, forthcoming.

[7] Anderson, Dennis, and John Claxton (1982). "Barriers to Consumer Choice of Energy Efficient Products." *Journal of Consumer Research*, Vol. 9, No. 2 (September), pages 163-170.

[8] Anderson, Soren, Ryan Kellogg, and James Sallee (2011). "What Do Consumers Believe about Future Gasoline Prices?" NBER Working Paper No. 16974 (April).

[9] Andreoni, James, and Charles Sprenger (2012). "Estimating Time Preferences from Convex Budgets." *American Economic Review*, Vol. 12, No. 7 (December), pages 3333-3356.

[10] 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.

[11] Baicker, Kate, Sendhil Mullainathan, and Joshua Schwartzstein (2012). "Choice Hazard in Health Insurance." Working Paper, Dartmouth College (October).

[12] Becker, Gary, and Kevin Murphy (1993). "A Simple Theory of Advertising as a Good or Bad." *Quarterly Journal of Economics*, Vol. 108, No. 4 (November), pages 941-964.

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

[14] 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.

[15] Bhargava, Saurabh, and Dayanand Manoli (2013). "Why Are Benefits Left on the Table? Assessing the Role of Information, Complexity, and Stigma on Take-up with an IRS Field Experiment." Working Paper, University of Texas at Austin.

[16] Blumstein, Carl, Betsy Krieg, Lee Schipper, and Carl York (1980). "Overcoming Social and Institutional Barriers to Energy Conservation." *Energy*, Vol. 5, pages 355-371.

[17] Borenstein, Severin, and Lucas Davis (2012). "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets." *Journal of Law and Economics*, Vol. 55, No. 1 (February).

[18] Bollinger, Brian, Phillip Leslie, and Alan Sorensen (2011). "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy*, Vol. 3, No. 1 (February), pages 91-128.

[19] Bureau of Labor Statistics (BLS) (2013). "Consumer Expenditure Survey: Expenditure Tables." http://www.bls.gov/cex/csxstnd.htm

[20] Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer (2013). “Are Consumers Myopic? Evidence from New and Used Car Purchases.” *American Economic Review*, Vol. 103, No. 1 (February), pages 220-256.

[21] Cabral, Marika, and Caroline Hoxby (2012). "The Hated Property Tax: Salience, Tax Rates, and Tax Revolts." NBER Working Paper 18514 (November).

[22] 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.

[23] Chetty, Raj (2009). "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics*, Vol. 1, pages 451-488.

[24] Chetty, Raj, Adam Looney, and Kory Kroft (2007). "Salience and Taxation: Theory and Evidence." NBER Working Paper 13330 (August).

[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] Choi, James, David Laibson, and Brigitte Madrian (2010). "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies*, Vol. 23, No. 4, pages 1405-1432.

[27] Choi, James, David Laibson, and Brigitte Madrian (2011). "$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans." *Review of Economics and Statistics*, Vol. 93, No. 3 (August), pages 748-763.

[28] Conlisk, John (1988). "Optimization Cost." *Journal of Economic Behavior and Organization*, Vol. 9, No. 3, pages 213-228.

[29] Davis, Lucas (2012). “Evaluating the Slow Adoption of Energy Efficient Investments: Are Renters Less Likely to Have Energy Efficient Appliances?” in "The Design and Implementation of U.S. Climate Policy," Edited by Don Fullerton and Catherine Wolfram. University of Chicago Press.

[30] Davis, Lucas, and Erich Muehlegger (2010). "Do Americans Consume too Little Natural Gas? An Empirical Test of Marginal Cost Pricing." *RAND Journal of Economics*, Vol. 41, No. 4 (Winter), pages 791-810.

[31] Deutsch, Matthias (2010a). "Life Cycle Cost Disclosure, Consumer Behavior, and Business Implications: Evidence from an Online Field Experiment." *Journal of Industrial Ecology*, Vol. 14, No. 1, pages 103-120.

[32] Deutsch, Matthias (2010b). "The Effect of Life-Cycle Cost Disclosure on Consumer Behavior: Evidence from a Field Experiment with Cooling Appliances." *Energy Efficiency*, Vol. 3, pages 303-315.

[33] DEHWA (Australian Department of the Environment, Water, Heritage, and the Arts) (2008). "Regulatory Impact Statement: Proposal to Phase-Out Inefficient Incandescent Light Bulbs." http://www.energyrating.gov.au/wp-content/uploads/Energy_Rating_Documents/Library/Lighting/Incandescent_Lamps/200808-ris-phaseout.pdf

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

[35] DOE (U.S. Department of Energy) (2009). "Technical Support Document: Impacts on the Nation of the Energy Independence and Security Act of 2007." http://www1.eere.energy.gov/buildings/appliance_standards/pdfs/en_masse_tsd_march_2009.pdf

[36] DOE (U.S. Department of Energy) (2010). "Energy Star CFL Market Profile: Data Trends and Market Insights." http://www.energystar.gov/ia/products/downloads/CFL_Market_Profile_2010.pdf,

[37] DOE (U.S. Department of Energy) (2012). "2010 US Lighting Market Characterization." http://apps1.eere.energy.gov/buildings/publications/pdfs/ssl/2010-lmc-final-jan-2012.pdf

[38] Dranove, David, and Ginger Zhe Jin (2010). "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature*, Vol. 48, No. 4 (December), pages 935-963.

[39] Duarte, Fabian, and Justine Hastings (2012). "Fettered Consumers and Sophisticated Firms: Evidence from Mexico's Privatized Social Security Market." NBER Working Paper 18582 (November).

[40] 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.

[41] Duflo, Esther, and Emmanuel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, Vol. 118, No. 3 (August), pages 815-842.

[42] Dupas, Pascaline (2011), "Do Teenagers Respond to HIV Risk Information? Evidence From a Field Experiment in Kenya", American Economic Journal: Applied Economics, Vol. 3, No. (January), pages 1-34.

[43] Energy Star (2013). "ENERGY STAR Unit Shipment and Market Penetration Report Calendar Year 2012 Summary." http://www.energystar.gov/ia/partners/downloads/unit_shipment_data/2012_USD_Summary_Rep 1224

[44] Falk, Armin and Florian Zimerrmann (2012). "A Taste for Consistency and Survey Response Behavior." *Cesifo Economic Studies*, Vol 59, No. 1, pages 181-193.

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

[46] Figlio, David, and Maurice E. Lucas (2004). "What's in a Grade? School Report Cards and the Housing Market." *American Economic Review*, Vol. 94, No. 3, pages 591–604.

[47] 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.

[48] Fong, Christina, and Erzo F.P. Luttmer (2009). "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty." *American Economic Journal: Applied Economics*, Vol. 1, No. 2 (April), pages 64-87.

[49] Formisano, Bob (2008). "2007 Energy Bill - Are They Phasing Out or Making Incandescent Bulbs Illegal?" http://homerepair.about.com/od/electricalrepair/ss/2007_energybill.htm

[50] Gabaix, Xavier (2013). "A Sparsity-Based Model of Bounded Rationality." Working Paper, New York University (July).

[51] 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.

[52] Gillingham, Kenneth, Matthew Harding, and David Rapson (2012). "Split Incentives and Household Energy Consumption." *Energy Journal*, Vol. 33, No. 2, pages 37-62.

[53] Gillingham, Kenneth, and Karen Palmer (2013). "Bridging the Energy Efficiency Gap." Resources for the Future Discussion Paper 13-02.

[54] Goldberg, Pinelopi (1998). "The Effects of the Corporate Average Fuel Economy Standards in the US." *Journal of Industrial Economics*, Vol. 46, pages 1-33.

[55] Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen (2006). "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments." *Quarterly Journal of Economics*, Vol. 121, No. 2, pages 399–460.

[56] Gruber, Jonathan, and Botond Koszegi (2004). "Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes." *Journal of Public Economics*, Vol. 88, pages 1959-1987.

[57] Gul, Faruk, and Wolfgang Pesendorfer (2007). "Harmful Addiction." *Review of Economic Studies*, Vol. 74, No. 1 (January), pages 147-172.

[58] Hassett, Kevin, and Gilbert Metcalf (1995). "Energy Tax Credits and Residential Conservation Investment: Evidence from Panel Data." *Journal of Public Economics*, Vol. 57, pages 201-217.

[59] Hastings, Justine, and Jeffrey Weinstein (2008).“Information, School Choice, and Academic Achievement: Evidence from Two Experiments.” *Quarterly Journal of Economics*, Vol. 123, No. 4, pages 1373–1414.

[60] 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.

[61] Heiss, Florian, Daniel McFadden, and Joachim Winter (2007). "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans." NBER Working Paper 13627.

[62] Herberich, David, John List, and Michael Price (2011). “How Many Economists Does it Take to Change a Light Bulb? A Natural Field Experiment on Technology Adoption.” Working Paper, University of Chicago (March).

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

[64] 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).

[65] 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.

[66] Houde, Sebastien (2012). "How Consumers Respond to Certification: A Welfare Analysis of the Energy Star Program." Working Paper, Stanford University (October).

[67] Interagency Working Group on Social Cost of Carbon, United States Government (2013). "Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866." http://www.whitehouse.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update.pdf

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

[69] Jensen, Robert (2010). “The Perceived Returns to Education and the Demand for Schooling.” *Quarterly Journal of Economics*, Vol. 125, No. 2 (May), pages 515-548.

[70] Jin, Ginger Zhe, and Alan Sorensen (2006). “Information and Consumer Choice: The Value of Publicized Health Plan Ratings.” *Journal of Health Economics*, Vol. 25, No. 2, pages 248–275.

[71] Kallbekken, Steffen, Hakon Saelen, and Erlend Hermansen (2013). "Bridging the Energy Efficiency Gap: A Field Experiment on Lifetime Energy Costs and Household Appliances." *Journal of Consumer Policy*, Vol. 36, pages 1-16.

[72] Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel (2012). "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics*. Vol. 127, No. 1, pages 199-235.

[73] Koszegi, Botond, and Adam Szeidl (2013). "A Model of Focusing in Economic Choice." *Quarterly Journal of Economics*, Vol. 128, No. 1, pages 53-107.

[74] Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, Vol. 112, No. 2 (May), pages 443-477.

[75] Levitt, Steve, and John List (2009). "Field Experiments in Economics: The Past, the Present, and the Future." European Economic Review, Vol. 53, pages 1-18.

[76] Luo, Renfu, Yaojian Shi, Linxiu Zhang, Huiping Zhang, Grant Miller, Alexis Medina, and Scott Rozelle (2012). "The Limits of Health and Nutrition Education: Evidence from Three Randomized-Controlled Trials in Rural China." *CESifo Economic Studies*, Vol. 58, No. 2, pages 385-404.

[77] Metcalf, Gilbert, and Kevin Hassett. "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data." *Review of Economics and Statistics*, Vol. 81, No. 3 (August), pages 517-528.

[78] Mills, Bradford, and Joachim Schleich (2010). "Why Don't Households See the Light? Explaining the Diffusion of Compact Fluorescent Lamps." *Resource and Energy Economics*, Vol. 32, pages 363-378.

[79] NHTSA (National Highway Traffic Safety Administration) (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).

[80] Newell, Richard, and Juha Siikamaki (2013). "Nudging Energy Efficiency Behavior: The Role of Information Labels." Resources for the Future Discussion Paper 13-17 (July).

[81] NRDC (Natural Resources Defense Council) 2011. "Better Light Bulbs Equal Consumer Savings in Every State." www.nrdc.org/policy

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

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

[84] Pope, Devin (2009). "Reacting to Rankings: Evidence from 'America's Best Hospitals.'" *Journal of Health Economics*, Vol. 28, No. 6, pages 1154-1165.

[85] Rabin, Matthew, and Georg Weizsacker (2009). "Narrow Bracketing and Dominated Choices." *American Economic Review*, Vol. 99, No. 4 (September), pages 1508-1543.

[86] Reis, Ricardo (2006). "Inattentive Consumers." *Journal of Monetary Economics*, Vol. 53, No. 8, pages 1761-1800.

[87] Sallee, James (2013). "Rational Inattention and Energy Efficiency." Working Paper, University of Chicago (September).

[88] Sallee, James, Sarah West, and Wei Fan (2009). "Consumer Valuation of Fuel Economy: A Microdata Approach." Working Paper, University of Chicago (September). Available from http://www.ntanet.org/images/stories/pdf/proceedings/09/032.pdf.

[89] Sanstad, Alan, and Richard Howarth (1994). "'Normal' Markets, Market Imperfections, and Energy Efficiency." *Energy Policy*, Vol. 22, No. 10, pages 811-818.

[90] Scanlon, Dennis, Michael Chernew, Catherine McLaughlin, and Gary Solon (2002). "The Impact of Health Plan Report Cards on Managed Care Enrollment." *Journal of Health Economics*, Vol. 21 No. 1, pages 19-41.

[91] Snyder, Mark (1974). "Self-Monitoring of Expressive Behavior." *Journal of Personality and Social Psychology*, Vol. 30, pages 526-537.

[92] Sylvania (2012). "5th Annual Sylvania Socket Survey." https://www.sylvania.com/en-us/tools-and-resources/surveys/Pages/socket-survey.aspx

[93] Sims, Christopher (2003). "Implications of Rational Inattention." *Journal of Monetary Economics*, Vol. 50, No. 3, pages 665-690.

[94] Snyder, Mark. (1974). "Self-monitoring of expressive behavior." *Journal of Personality and Social Psychology*, Vol. 30, pages 526–537.

[95] Strotz, Robert (1956). "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies*, Vol. 23, pages 165-180.

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

[97] Toledo, Chantal (2013). "Do Environmental Messages Work on the Poor? Experimental Evidence from Brazilian Favelas." Working Paper, University of California at Berkeley (January).

[98] Ward, David, Christopher Clark, Kimberly Jensen, Steven Yen, and Clifford Russell (2011). "Factors Influencing Willingness-to-Pay for the Energy Star Label." *Energy Policy*, Vol. 39, No. 3 (March), page 1450 – 1458.

Tables

Table 1: Descriptive Statistics and Balance for TESS Experiment

	(1)	(2)	(3)
Individual Characteristics	Population Mean	Treatment - Control Difference	Positive - Balanced Treatment Difference
Baseline Relative Willingness-to-Pay	2.9	0.20	-0.25
for CFL ($)	(7.1)	(0.52)	(0.70)
Household Income ($000s)	70.9	-2.86	-3.79
	(51.8)	(3.92)	(3.86)
Education (Years)	13.8	-0.04	0.18
	(2.5)	(0.18)	(0.21)
Age	46.7	0.26	0.22
	(16.9)	(1.26)	(1.45)
Male	0.48	-0.007	-0.009
	(0.50)	(0.035)	(0.040)
Liberal	0.00	0.056	-0.005
	(1.00)	(0.076)	(0.084)
Party	0.00	0.080	0.078
	(1.00)	(0.072)	(0.082)
Environmentalist	0.30	-0.024	0.019
	(0.32)	(0.023)	(0.026)
Conserve Energy	0.55	0.008	0.032
	(0.50)	(0.036)	(0.041)
Homeowner	0.70	0.022	-0.012
	(0.46)	(0.035)	(0.038)
F-Test p-Value		0.848	0.995

Notes: Column 1 presents means of individual characteristics in the TESS experiment population, with standard deviations in parenthesis. Column 2 presents differences in means between the treatment groups and control, while column 3 presents differences in means between Positive and Balanced treatment groups. Columns 2 and 3 have robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Table 2: Effects of TESS Informational Interventions

	(1)	(2)	(3)	(4)	(5)	(6)
1(Treatment)	2.535	2.301	2.324	2.078	3.231	2.138
	(0.549)***	(0.358)***	(0.364)***	(0.777)***	(0.364)***	(0.498)***
Baseline Willingness-to-Pay		0.777	0.775		0.934	0.776
		(0.037)***	(0.037)***		(0.065)***	(0.037)***
1(Positive Treatment)						0.396
						(0.573)
R2	0.03	0.56	0.57	0.02	0.33	0.57
N	1,203	1,203	1,188	656	919	1,188
Individual Characteristics	No	No	Yes	No	Yes	Yes
Include Endline-Only Group	No	No	No	Yes	No	No
Exclude Max./Min. Baseline WTP	No	No	No	No	Yes	No

Notes: This table presents estimates of Equation (1). The outcome variable is endline willingness-to-pay for the CFL. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Table 3: Perceived Intent of TESS Study

	(1)	(2)	(3)
Do you think that the intent of the study was to . . .	Control	Balanced Treatment	Positive Treatment
Understand the effect of price changes on purchasing patterns	0.44	0.34	0.37
Measure whether people make consistent purchases in similar situations	0.31	0.25	0.26
Understand why people buy incandescents vs. CFLs	0.31	0.48	0.47
Test how well people are able to quantify energy costs	0.27	0.38	0.46
Test whether ability to quantify energy costs affects purchases of incandescents vs. CFLs	0.33	0.50	0.54
Test whether the number of bulbs in a package affects purchasing patterns	0.37	0.22	0.26
Test whether consumer education affects purchases of incandescents vs. CFLs	0.41	0.60	0.64
Understand what features of lightbulbs are most important to people	0.30	0.41	0.34
Predict the future popularity of incandescents vs. CFLs	0.30	0.34	0.34
None of the above	0.05	0.08	0.05
Number of Respondents	461	545	519

Notes: This table presents the share of consumers in each group who responded that the intent of the study was as listed in the leftmost column. Observations are weighted for national representativeness.

Table 4: Effects on Beliefs

	(1)	(2)
	Savings Belief	Belief Error
1(Treatment)	13.0	-14.0
	(3.8)***	(1.8)***
Constant	25.0	34.0
	(3.8)***	(1.4)***
R2	.	.
N	1,506	1,506

Notes: In the post-experiment survey, consumers were asked their beliefs about the dollar value of electricity cost savings from owning CFLs instead of incandescents. Columns 1 and 2 present median regressions of the effects of the informational interventions on these beliefs and the absolute value of the error in these beliefs, respectively. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Table 5: Effects on Important Factors in Purchase Decision

	(1)	(2)	(3)	(4)	(5)
	Price	Energy Use	Bulb Lifetime	Warm-Up Time	Mercury and Disposal
1(Balanced Treatment)	-0.864	0.147	0.023	-0.943	-0.294
	(0.208)***	(0.214)	(0.201)	(0.243)***	(0.252)
1(Positive Treatment)	-0.552	0.202	0.249	0.036	-0.089
	(0.218)**	(0.210)	(0.187)	(0.231)	(0.244)
Constant	7.747	7.435	7.760	5.406	6.030
	(0.134)***	(0.145)***	(0.133)***	(0.167)***	(0.178)***
R2	0.02	0.00	0.00	0.02	0.00
N	1,533	1,478	1,512	1,506	1,518

Notes: This table reports treatment effects on self-reported importance of different factors in purchase decisions. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Table 6: Descriptive Statistics and Balance for In-Store Experiment

	(1)	(2)	(3)	(4)
Individual Characteristics	Experimental Sample Mean	Refused - Sample Difference	Treatment - Control Difference	Rebate - Standard Difference
Energy an Important Factor	0.25		0.009	-0.024
in Purchase Decision	(0.43)		(0.026)	(0.026)
Expected Usage (Minutes/Day)	333		12.8	2.7
	(280)		(17.0)	(17.0)
Age	43.8	2.3	0.7	-0.3
	(11.4)	(0.6)***	(0.7)	(0.7)
Male	0.66	0.06	0.009	0.003
	(0.47)	(0.03)**	(0.029)	(0.029)
African American	0.16	-0.04	-0.001	-0.008
	(0.37)	(0.02)**	(0.022)	(0.022)
Asian	0.06	0.04	-0.030	0.005
	(0.24)	(0.02)**	(0.014)**	(0.015)
Caucasian	0.66	-0.07	0.037	-0.005
	(0.47)	(0.03)**	(0.029)	(0.029)
Hispanic	0.07	0.06	0.001	0.011
	(0.25)	(0.02)***	(0.015)	(0.015)
Middle Eastern	0.01	0.01	0.002	0.007
	(0.12)	(0.01)	(0.013)	(0.007)
F-Test p-Value		0.00	0.742	0.896
Purchase Decisions				
Purchased Any Lightbulb	0.77		0.011	0.027
	(0.42)		(0.025)	(0.025)
Purchased Substitutable Lightbulb	0.73		-0.008	0.011
	(0.44)		(0.027)	(0.027)

Notes: Column 1 presents means of individual characteristics in the in-store experiment sample, with standard deviations in parenthesis. Column 2 presents differences in recorded demographic characteristics between those who refused or did not complete the survey and the experimental sample. Column 3 presents differences in means between treatment and control groups, while column 4 presents differences in means between the rebate and standard coupon groups. Columns 2, 3, and 4 have robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

Table 7: Effects of In-Store Informational Intervention

	(1)	(2)	(3)
1(Rebate)	0.094	0.105	0.078
	(0.035)***	(0.033)***	(0.047)*
1(Treatment)	-0.002	0.004	-0.022
	(0.035)	(0.033)	(0.045)
1(Rebate and Treatment)			0.054
			(0.066)
R2	0.01	0.16	0.16
N	794	793	793
Individual Characteristics	No	Yes	Yes

Notes: This table presents estimates of Equation (2), a linear probability model with outcome variable 1(Purchased CFL). Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

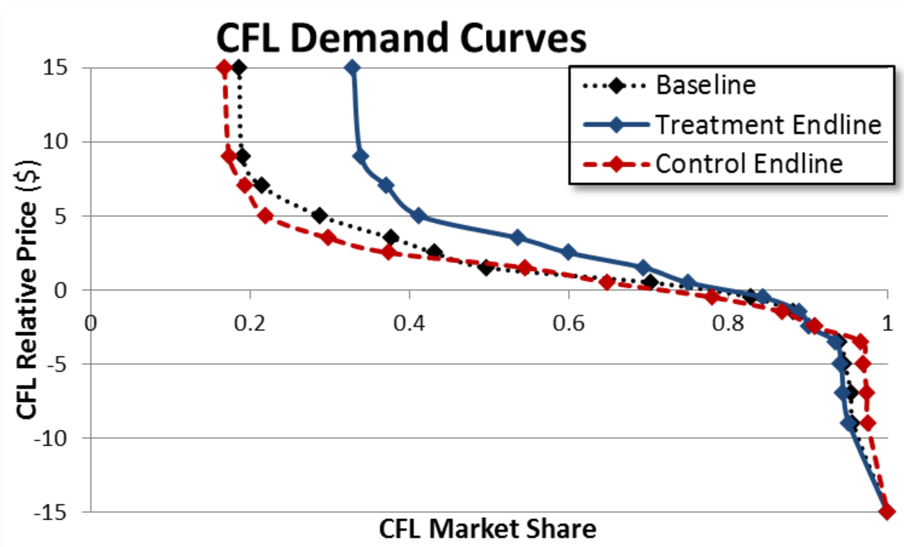
Table 8: Welfare Analysis Using TESS Results

(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy ($/package)	Relative WTP (v_i^0) of Marginal Consumers ($/package)	Average Marginal Internality ($/package)	Demand Density (Share of packages)	Marginal Welfare Effect ($/package)	Cumulative Welfare Effect ($/package)
1	0.5	2.11	0.126	0.204	0.204
2	1.5	2.16	0.052	0.034	0.238
3	2.5	3.41	0.028	0.026	0.264
4	3.5	1.77	0.030	-0.052	0.212
6	5	1.77	0.006	-0.020	0.192
8	7	1.77	0.008	-0.042	0.150
10	9	1.77	0.003	-0.019	0.131
∞	15	1.77	0.043	-0.567	-0.436

Notes: This table uses the TESS experiment results to calculate the welfare effects at different levels of the CFL subsidy. Observations are weighted for national representativeness. See text for details.

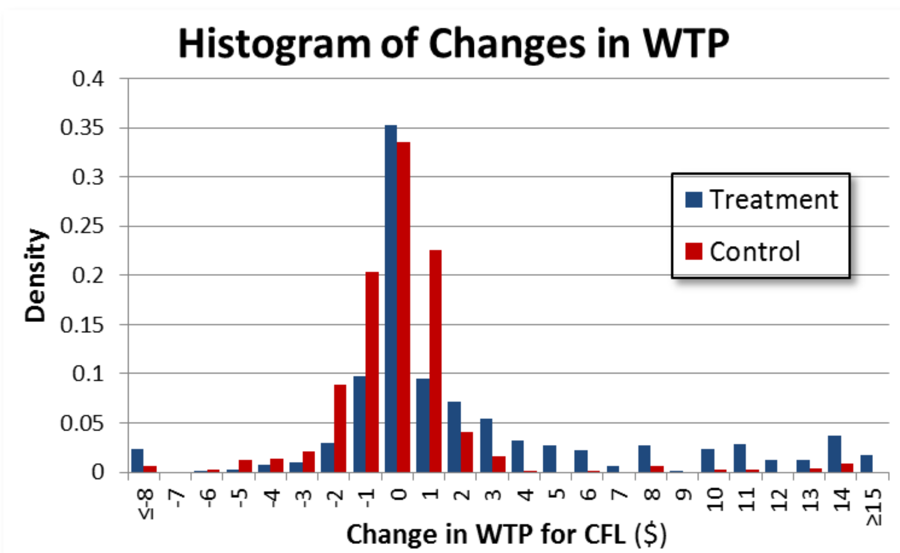
Figures

Figure 1: Demand Curves



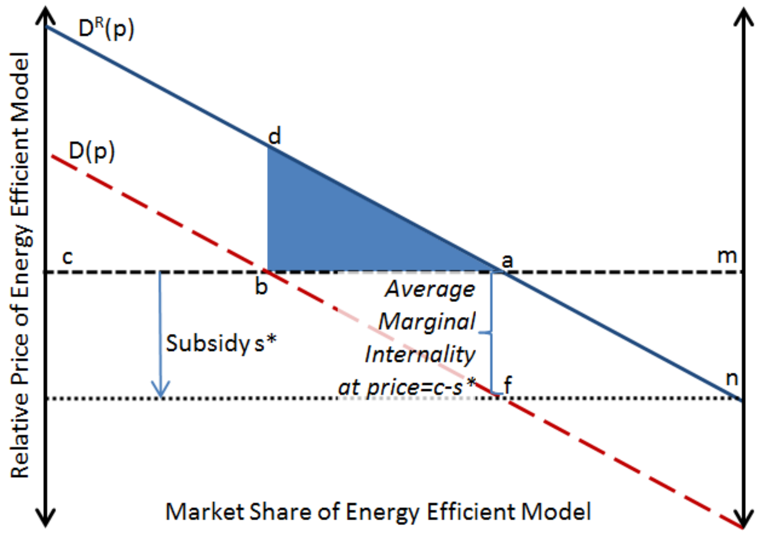
Note: This figure plots the baseline and endline demand curves from the TESS experiment. Observations are weighted for national representativeness.

Figure 2: Histogram of Changes in WTP



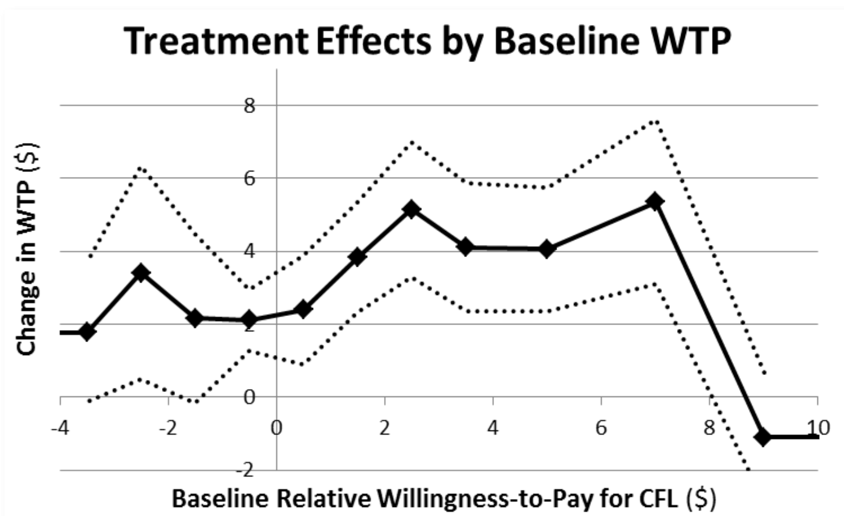
Note: This figure plots the histogram of changes in willingness-to-pay for the CFL for the treatment and control groups. Observations are weighted for national representativeness.

Figure 3: Theoretical Framework



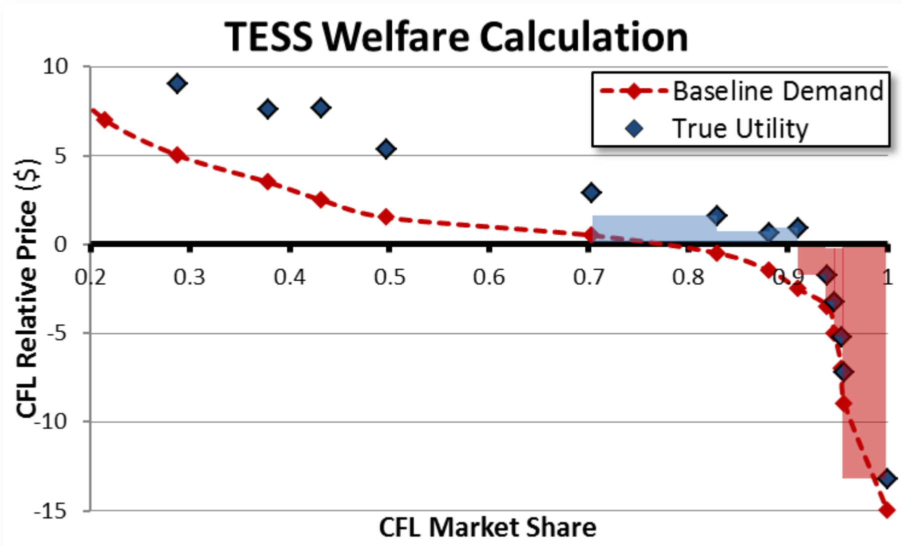
Notes: This figure illustrates the theoretical framework. $D(p)$ is the market demand curve. $D^R(p)$ is the demand curve if all consumers are unbiased. The dashed horizontal line represents c, the relative cost of the energy efficient CFL, while the dotted horizontal line is the after-subsidy relative price.

Figure 4: TESS Experiment Treatment Effects by Initial WTP



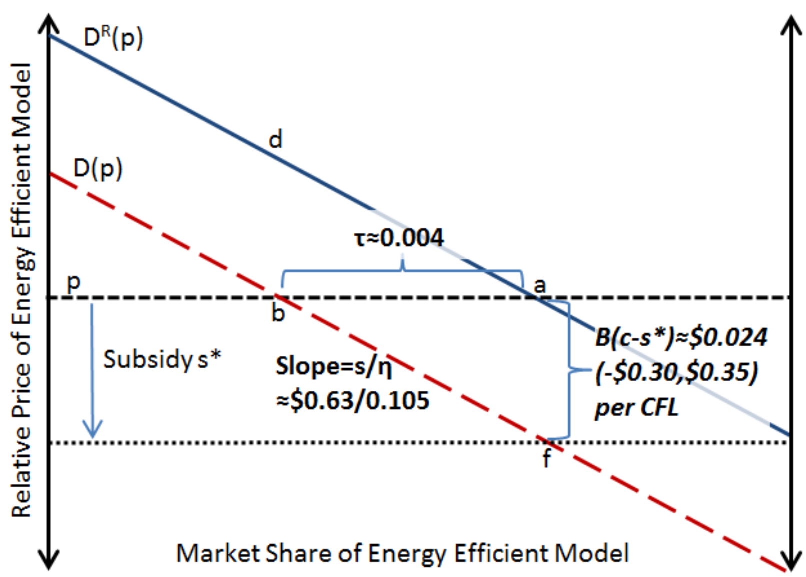
Note: This graph presents the average treatment effects of the informational intervention for each level of baseline relative willingness-to-pay. Due to limited sample size, baseline WTPs of less than -$3.50 are collapsed to -$3.50, and baseline WTPs of $9 and $15 are collapsed to $9. Dotted lines are 90 percent confidence intervals. Observations are weighted for national representativeness.

Figure 5: Welfare Calculation Using TESS Experiment



Notes: This figure illustrates the welfare effects of marginal increases in the CFL subsidy using the TESS experiment results. Observations are weighted for national representativeness.

Figure 6: Policy Analysis Using In-Store Experiment



Notes: This figure illustrates the use of treatment effects of rebates and information as sufficient statistics for a first-order approximation to the optimal subsidy. Coefficient estimates are from the in-store experiment.

Online Appendix: Not For Publication

The Lightbulb Paradox: Evidence from Two Randomized Experiments

Hunt Allcott and Dmitry Taubinsky

Appendix 1: Details of TESS Experiment

Introductory Screen

In appreciation for your participation in this study, we are giving you a $10 shopping budget. During the study, we will ask you to make 30 decisions between pairs of light bulbs using that $10 shopping budget. There will be a first set of 15 decisions, then a break, and then a second set of 15 decisions.

After you finish with all 30 decisions, one of them will be randomly selected as your "official purchase." In approximately four to six weeks, GfK will send you the light bulbs you chose in that official purchase. After your official purchase has been paid for from the $10 shopping budget we are giving you, any money left over will be provided to you in the form of **bonus points awarded to your account.** This means that after the study is completed, you will receive 1) the light bulbs you selected in the decision that is randomly selected to be your "official purchase" and 2) an amount between zero and 10000 bonus points, corresponding to whatever money is left in your shopping budget after the purchase.

Light bulbs are frequently shipped in the mail. There is not much risk of breakage, but if anything does happen, GfK will just ship you a replacement. Even if you don't need light bulbs right now, remember that you can store them and use them in the future.

Since each of your decisions has a chance of being your official purchase, you should think about each decision carefully.

Next

Baseline Lightbulb Choices (Top of Screen)

We have given you a $10 shopping budget to purchase a package of light bulbs. Your first 15 purchase decisions will concern the two packages of light bulbs shown below.

Choice A	Choice B
Philips 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	**Philips 60-Watt Incandescent Light Bulbs, 4-Pack**
	

Click for detailed product information

Between the 15 decisions, the only thing that varies is the price. Each of these decisions has a chance of being the one choice (out of 30) that will become your official purchase, so you should think about each purchase carefully. Whatever money you do not spend on the light bulbs, you get to keep: any remaining money will be provided to you as cash-equivalent bonus points. Please think about each decision carefully.

Here is an example of how this might work. After you make all your decisions, suppose that Decision Number 6 from the set below were selected as your official purchase.

- If you had chosen Choice A, you would pay $2 from your $10 shopping budget. You would receive the Choice A light bulb package in the mail within 4-6 weeks, as well as the remaining $10-$2=$8 in your shopping budget (You would receive that $8 in the form of 8000 bonus points credited to your account.)
- If you had chosen Choice B, you would pay $4 from your $10 shopping budget. You would receive the Choice B light bulb package in the mail within 4-6 weeks, as well as the remaining $10-$4=$6 in your shopping budget. (You would receive that $6 in the form of 6000 bonus points credited to your account.)

Now please make your decisions for each of the 15 choices below.

Detailed Product Information Screen

Detailed Product Information

Choice:	A	B
Manufacturer:	Philips	Philips
Type:	Compact Fluorescent (CFL)	Incandescent
Number of Bulbs:	1	4
Light Output:	60 Watt-equivalent	60 Watts
Light Output:	900 Lumens	840 Lumens
Color Temperature:	2700K	2700K
Energy Use:	13 Watts	60 Watts
Manufacturer's Home Country:	USA	USA

Back to Questionnaire

Baseline Lightbulb Choices (Bottom of Screen)

Now please make your decisions for each of the 15 choices below.

Decision Number	Choice A 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	Choice B 60-Watt Incandescent Light Bulbs, 4-Pack
	Purchase **Choice A for free**	Purchase **Choice B for $10**
1)	○	○
	Purchase **Choice A for free**	Purchase **Choice B for $8**
2)	○	○
	Purchase **Choice A for free**	Purchase **Choice B for $6**
3)	○	○
	Purchase **Choice A for free**	Purchase **Choice B for $4**
4)	○	○
	Purchase **Choice A for $1**	Purchase **Choice B for $4**
5)	○	○
	Purchase **Choice A for $2**	Purchase **Choice B for $4**
6)	○	○
	Purchase **Choice A for $3**	Purchase **Choice B for $4**
7)	○	○
	Purchase **Choice A for $4**	Purchase **Choice B for $4**
8)	○	○

Note: This does not show all of the 15 Decision Numbers.

Balanced Treatment Introductory Screen

For this next part of the study, you will have the opportunity to learn more about compact fluorescent light bulbs (CFLs) and incandescent light bulbs. We will focus on the following two issues:

- Total Costs
- Disposal and Warm-Up Time

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

Next

Total Cost Information Screen

CFLs last longer than incandescents. At average usage:

- Incandescents burn out and have to be replaced every year.
- CFLs burn out and have to be replaced every eight years.

If one incandescent bulb costs $1 and one CFL costs $4, this means that the total purchase prices for eight years of light are:

- $8 for incandescents
- $4 for CFLs

Also, CFLs use less electricity than incandescents. At national average usage and electricity prices:

- A standard (60-Watt) incandescent uses $6 in electricity each year.
- An equivalent CFL uses $1.50 in electricity each year.

Thus, for eight years of light, the total costs to purchase bulbs and electricity would be:

- $56 for incandescents: $8 for the bulbs plus $48 for electricity
- $16 for a CFL: $4 for the bulbs plus $12 for electricity

The graph below illustrates this:

Question: For eight years of light, how much larger are the total costs (for bulbs plus electricity) for 60-Watt incandescents as compared to their CFL equivalents?

To answer this question, please enter a dollar amount in the first number box and the cent amount in the second number box. For example, if you wanted to enter X dollars and Y cents, you would enter X in the first number box and Y in the second number box.

Type your answer below.

$ [ ] . [ ]

Disposal and Warm-Up Information Screen

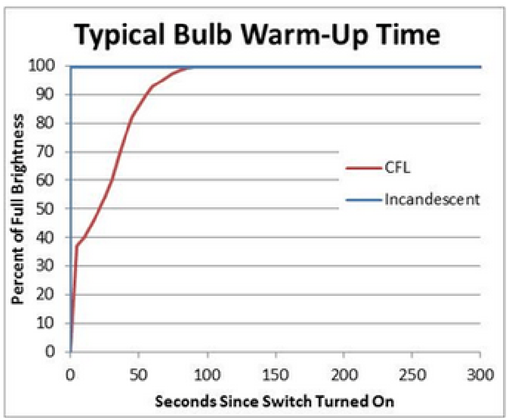
After they burn out, CFLs need proper disposal:

- Because CFLs contain mercury, it is recommended that they be properly recycled, and not simply disposed of in regular household trash. CFLs can be recycled through:
  - Local waste collection sites
  - Mail-back services that you can find online
  - Local retailers, including Ace Hardware, IKEA, Home Depot, and Lowe's, as well as other retailers.
- No special precautions need to be taken to dispose of an incandescent light bulb. Incandescents can be disposed of in regular household trash.

After the light switch is turned on, CFLs take longer to warm up than incandescents:

- An incandescent reaches full brightness immediately.
- A typical CFL can take 60 to 90 seconds to reach its full brightness.

The graph below illustrates this:



<u>Question</u>: About how much longer does it take a typical CFL to reach full brightness, as compared to an incandescent?

Type your answer below.

☐ to ☐ seconds

Control Introductory Screen

For this next part of the study, you will have the opportunity to learn more about light bulbs. We will focus on the following two issues:

1. Number of Bulbs by Sector
2. Sales Trends

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

Next

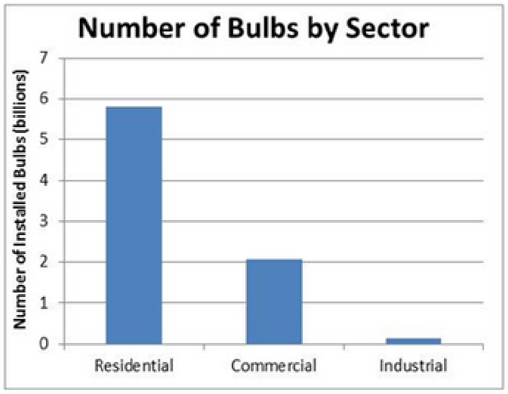
Number of Bulbs by Sector Information Screen

According to official estimates, there are slightly more than eight billion light bulbs installed in the United States.

The US economy can be divided into three major sectors: residential, commercial, and industrial. Each sector has a different number of light bulbs:

- There are about 5.8 billion light bulbs installed in residential buildings in the U.S.
- There are about 2.1 billion light bulbs installed in commercial buildings in the U.S.
- There are about 0.14 billion light bulbs installed in industrial buildings in the U.S.

The graph below illustrates this:



Question: About how many more light bulbs are installed in residential buildings compared to commercial buildings in the U.S.?

To answer this question, you can enter whole numbers and/or decimals.

Type your answer below.

billion

Sales Trends Information Screen

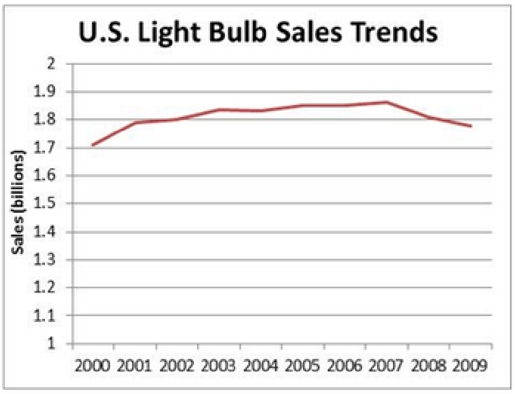
According to official sales data, sales of light bulbs in the United States have had the following trend:

- Sales increased in each year between 2000 and 2007.
- Sales decreased slightly in 2008 and 2009.

Total light bulb sales were different at the end of the decade compared to the beginning:

- Sales in 2000 were just over 1.7 billion bulbs.
- Sales in 2009 were just under 1.8 billion bulbs.

The graph below illustrates this:



Question: About how many light bulbs were sold in the United States in 2009?

To answer this question, you can enter whole numbers and/or decimals.

Type your answer below.

[ ] billion

Endline Lightbulb Choices (Top of Screen)

Remember, we have given you a $10 shopping budget to purchase a package of light bulbs. Your second 15 purchase decisions will concern the two packages of light bulbs shown below.

Choice A
Philips 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack

Choice B
Philips 60-Watt Incandescent Light Bulbs, 4-Pack

Click for detailed product information

Between the 15 decisions, the only thing that varies is the price. Each of these decisions has a chance of being the one choice (out of 30) that will become your official purchase, so you should think about each purchase carefully. Whatever money you do not spend on the light bulbs, you get to keep: any remaining money will be provided to you as cash-equivalent bonus points. Please think about each decision carefully.

Here is an example of how this might work. After you make all your decisions, suppose that Decision Number 21 from the set below were selected as your official purchase.

- If you had chosen Choice A, you would pay $2 from your $10 shopping budget. You would receive the Choice A light bulb package in the mail within 4-6 weeks, as well as the remaining $10-$2=$8 in your shopping budget (You would receive that $8 in the form of 8000 bonus points credited to your account)
- If you had chosen Choice B, you would pay $4 from your $10 shopping budget. You would receive the Choice B light bulb package in the mail within 4-6 weeks, as well as the remaining $10-$4=$6 in your shopping budget (You would receive that $6 in the form of 6000 bonus points credited to your account.)

Now please make your decisions for each of the 15 choices below.

Endline Lightbulb Choices (Bottom of Screen)

Now please make your decisions for each of the 15 choices below.

Decision Number	Choice A 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	Choice B 60-Watt Incandescent Light Bulbs, 4-Pack
16)	Purchase **Choice A for free**	Purchase **Choice B for $10**
17)	Purchase **Choice A for free**	Purchase **Choice B for $8**
18)	Purchase **Choice A for free**	Purchase **Choice B for $6**
19)	Purchase **Choice A for free**	Purchase **Choice B for $4**
20)	Purchase **Choice A for $1**	Purchase **Choice B for $4**
21)	Purchase **Choice A for $2**	Purchase **Choice B for $4**
22)	Purchase **Choice A for $3**	Purchase **Choice B for $4**
23)	Purchase **Choice A for $4**	Purchase **Choice B for $4**
	Purchase **Choice A for $4**	Purchase **Choice B for $3**

Note: This does not show all of the 15 Decision Numbers.

Post-Experiment Survey Questions

Question 1. *How important were the following factors in your purchase decision? [Rate from 1-10]*

1. *Energy use*
2. *Time required for the bulb to reach full brightness after it is turned on*
3. *Bulb lifetime*
4. *Mercury content and protocols for proper disposal*
5. *Purchase Ppice*

Question 2. *Do you think that the intent of the study was to ...*
Select all answers that apply

1. *Understand the effect of price changes on purchasing patterns*
2. *Measure whether people make consistent purchases in similar situations*
3. *Understand why people buy incandescents vs. CFLs*
4. *Test how well people are able to quantify energy costs*
5. *Test whether ability to quantify energy costs affects purchases of incandescents vs. CFLs*
6. *Test whether the number of bulbs in a package affects purchasing patterns*
7. *Test whether consumer education affects purchases of incandescents vs. CFLs*
8. *Understand what features of lightbulbs are most important to people*
9. *Predict the future popularity of incandescents vs. CFLs*
10. *None of the above*

Question 3. Part A: *The typical CFL lasts 8000 hours, or about eight years at typical usage rates. Do you think it costs more or less to buy electricity for that 8000 hours of light from compact fluorescent light bulbs (CFLs) compared to incandescent light bulbs?*

- More
- Less

Part B: *At national average electricity prices, how much [more/less] does it cost to buy electricity for that 8000 hours of light from compact fluorescent light bulbs (CFLs) compared to incandescent light bulbs? Just give your best guess.*

Question 4. *Some states and local areas have rebates, low-interest loans, or other incentives available for energy efficiency. These might include rebates for Energy Star appliances or energy efficient light bulbs, low-interest loans for energy-saving home improvements, government-funded weatherization, and other programs. Are any such programs available in your area?*

1. *Yes*
2. *I think so, but I'm not sure*
3. *I'm not sure at all*
4. *I think not, but I'm not sure*
5. *No*

Question 5. *This question is about hypothetical choices and does not affect your earnings in this study.*

Suppose that you could get the amount under "Option A" (i.e. $100), or the amount under "Option B" a year later. Assume it's no more work for you to receive the money under Option A than under Option B, and that you would receive the money for sure, regardless of when you choose to receive it. Which would you prefer?

For each row, please choose whether you would prefer the payment under "Option A" or under "Option B".

	Option A	Option B
1)	$100 today	$50 in one year
2)	$100 today	$90 in one year
3)	$100 today	$100 in one year
4)	$100 today	$110 in one year
5)	$100 today	$130 in one year
6)	$100 today	$150 in one year
7)	$100 today	$170 in one year
8)	$100 today	$200 in one year
9)	$100 today	$250 in one year
10)	$100 in one year	$50 in two years
11)	$100 in one year	$90 in two years
12)	$100 in one year	$100 in two years

Notes: This does not show all of the 18 choices. Participants were randomly assigned to receive either this table or another table that was identical except that the bottom half and top half were switched, so that the one year vs. two year tradeoffs were presented first.

Question 6. *Please indicate how much you agree or disagree with the following statements:*
Select one answer from each row in the grid
[Strongly Agree Agree Neutral Disagree Strongly Disagree]

1. *It's important to me to fit in with the group I'm with.*
2. *My behavior often depends on how I feel others wish me to behave.*
3. *My powers of intuition are quite good when it comes to understanding others' emotions and motives.*
4. *My behavior is usually an expression of my true inner feelings, attitudes, and beliefs.*
5. *Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.*
6. *I would NOT change my opinions (or the way I do things) in order to please someone else or win their favor.*

Appendix 2: Additional TESS Results

Baseline Willingness-to-Pay

Table A2-1 shows the association between baseline WTP v^0 and a series of individual characteristics. Column 1 shows that men, democrats, and those who report having taken steps to conserve energy have higher demand for CFLs. Columns 2-5 separately test individual variables of environmentalism and political ideology which are correlated, providing additional evidence that liberals tend to have higher WTP. These correlations conform to our intuition and build further confidence that the differences in WTP are meaningful.

The table also provides suggestive evidence on two distortions other than imperfect information and inattention which might justify subsidies and standards. The first is a particular form of agency problem in real estate markets: renters might have lower CFL demand because they might leave the CFLs in the house's light sockets when they move and be unable to capitalize on their investment. Lacking random or quasi-random assignment in renter vs. homeowner status, Davis (2012) and Gillingham, Harding, and Rapson (2012) correlate durable good ownership with homeowner status conditional on observables. Column 1 replicates their approach in the TESS data, showing no conditional association between WTP and homeowner status. Column 6 shows that the unconditional association is also statistically zero.

The second potential distortion considered in Table A2-1 is present bias. In the post-experiment survey, we estimate the β and δ of a quasi-hyperbolic model through a menu of hypothetical intertemporal choices at two different time horizons: \$100 now vs. $\$m_i^1$ in one year, and \$100 in one year vs. $\$m_i^2$ in two years. Denoting $\widehat{m}^1$ and $\widehat{m}^2$ as the minimum values at which participant i prefers money sooner, the long run discount factor is $\delta_i = 100/\widehat{m}^2$, and the present bias parameter is $\beta_i = \widehat{m}_i^2/\widehat{m}_i^1$. We dropped non-monotonic responses and top-coded $\widehat{m}^1$ and $\widehat{m}^2$ analogously to how we constructed v^0 and v^1.

If there is a distribution of β and δ, consumers with higher β and δ should be more likely to purchase CFLs. Column 1 shows that there is a conditional correlation between δ and baseline WTP v_i^0, suggesting that people who are more patient may be more likely to purchase CFLs. However, there is no statistically significant correlation between β and v_i^0. Column 7 repeats the estimates without any conditioning variables, and the coefficients are comparable. The results in column 1 rule out with 90 percent confidence that a one standard deviation increase in β increases WTP for the CFL by more than \$0.47. In sum, these correlations provide no suggestive evidence in favor of the hypotheses that agency problems or present bias play a role in lightbulb decisions.

Table A2-1: Association Between Individual Characteristics and CFL Demand

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income (000s)	0.005						
	(0.006)						
Education (Years)	0.010						
	(0.148)						
Age	0.003						
	(0.018)						
Male	0.931						
	(0.533)*						
Liberal	0.091	0.374					
	(0.389)	(0.285)					
Party	0.573		0.562				
	(0.344)*		(0.266)**				
Environmentalist	0.682			1.429			
	(0.791)			(0.804)*			
Conserve Energy	0.970				0.863		
	(0.525)*				(0.545)		
Homeowner	0.047					0.116	
	(0.716)					(0.616)	
Present Bias β	0.281						0.144
	(0.298)						(0.284)
Discount Factor δ	1.215						0.962
	(0.620)*						(0.605)
R2	0.03	0.00	0.01	0.00	0.00	0.00	0.01
N	1,163	1,226	1,229	1,221	1,219	1,229	1,178

Notes: Left-hand-side variable is baseline relative WTP for the CFL. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Table A2-2: Correlation of Treatment Effects with Self-Monitoring Scale

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Important to fit in	0.206						
	(0.399)						
Behave as others wish		0.480					
		(0.384)					
Good intuition for others' motives			0.266				
			(0.295)				
(-1)*Behavior expresses true feelings				-0.410			
				(0.332)			
Regulate my actions					-0.218		
					(0.310)		
(-1)*NOT change opinions to please someone						-0.114	
						(0.365)	
Self-Monitoring Mean							-0.010
							(0.008)
R2	0.57	0.57	0.56	0.57	0.56	0.56	0.57
N	1,185	1,184	1,184	1,184	1,184	1,184	1,188

Notes: This table presents estimates of Equation (1) with the addition of Self-Monitoring Scale variables and the interaction of these variables with the treatment indicator. The outcome variable is endline willingness-to-pay for the CFL. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Tables A2-3A and A2-3B: Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values

Assuming Top-Coded and Bottom-Coded WTPs Average \$12 and -\$12, Respectively

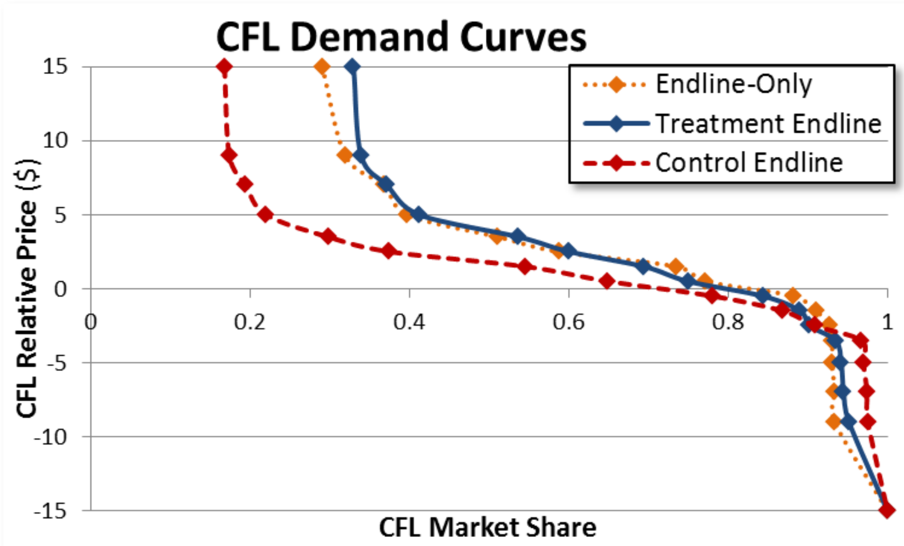
(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy (\$/package)	Relative WTP (v_i^0) of Marginal Consumers (\$/package)	Average Marginal Internality (\$/package)	Demand Density (Share of packages)	Marginal Welfare Effect (\$/package)	Cumulative Welfare Effect (\$/package)
1	0.5	1.92	0.126	0.180	0.180
2	1.5	2.29	0.052	0.041	0.221
3	2.5	2.98	0.028	0.014	0.234
4	3.5	0.69	0.030	-0.085	0.150
6	5	0.69	0.006	-0.027	0.123
8	7	0.69	0.008	-0.050	0.073
10	9	0.69	0.003	-0.022	0.051
∞	12	0.69	0.043	-0.485	-0.434

Assuming Top-Coded and Bottom-Coded WTPs Average \$20 and -\$20, Respectively

(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy (\$/package)	Relative WTP (v_i^0) of Marginal Consumers (\$/package)	Average Marginal Internality (\$/package)	Demand Density (Share of packages)	Marginal Welfare Effect (\$/package)	Cumulative Welfare Effect (\$/package)
1	0.5	2.43	0.126	0.244	0.244
2	1.5	1.95	0.052	0.024	0.268
3	2.5	4.11	0.028	0.046	0.314
4	3.5	0.18	0.030	-0.100	0.213
6	5	0.18	0.006	-0.030	0.183
8	7	0.18	0.008	-0.054	0.129
10	9	0.18	0.003	-0.023	0.106
∞	20	0.18	0.043	-0.850	-0.744

Notes: These tables use the TESS experiment results to calculate the welfare effects at different levels of the CFL subsidy. They replicate Table 8, except with different assumed mean censored values of WTP. Observations are weighted for national representativeness.

Figure A2-1: Post-Only Treatment Demand Curve



Notes: This presents the demand curve for the Endline-Only treatment group, along with demand curve for the control group and other treatment group consumers. Observations are weighted for national representativeness.

Appendix 3: iPad Total Cost Comparison Screen

Notes: This is the information screen presented to treatment group consumers in the in-store experiment. Numbers in this screen shot represent a consumer buying one CFL at typical purchase prices and national average electricity prices.

Appendix 4: Appendix to Theoretical Framework

Proof of Proposition 1

Rewrite welfare as

$$W(s) = C + \sum_k \alpha_k \left[\int \sigma(v, b_k, c-s)(v-c)dF_{v|b_k}(v|b_k) \right]$$
$$= C + \sum_k \alpha_k \left[\int_{v \geq c-s+b_k} (v-c)dF_{v|b_k}(v|b_k) \right]. \tag{12}$$

Differentiating (12) with respect to s yields

$$W'(s) = \sum_k \alpha_k \left[-((c-s+b_k)-c) f_{v|b_k}(c-s+b_k|b_k) \right]$$
$$= \sum_k \alpha_k (s-b_k) f_{v|b}(c-s+b_k|b_k)$$
$$= \sum_k (b_k - s) D'_{b_k}(c-s)$$
$$= (B(c-s)-s)\, D'(c-s).$$

Proof of Proposition 2

Suppose that consumers are homogeneous in their bias: $b_k \equiv b$ for all k. Then by Proposition 1, an optimal subsidy must satisfy $s^* = b(c-s^*)$. We now show that this subsidy attains the first best under homogeneity. Plugging $s^* = b(c-s^*)$ into the the social welfare function yields

$$W(s^*) = C + \int_{v \geq c-s^*+b(p-s^*)} (v-c)dF(v)$$
$$= C + \int_{v \geq c} (v-c)dF(v). \tag{13}$$

But by definition, (13) is just the first best level of welfare.

Set $\psi(s) = s - b(c-s)$. We must now show that $\psi(s) = 0$ has a solution s^*. By definition and by Assumption 1, $\psi'(s) = 1 + b'(c-s) \geq 1 + \rho$, where $\rho > -1$. This ensures that $\psi(s)$ has a unique solution.

When consumers are heterogenous in the way specified in the proposition, the argument in the body of the paper shows why the first best can't be obtained.

Generalizing the Analysis to Partial Internality Reduction

Suppose that consumer bias is given by $b = b^x + b^y$, and set

$$B^x(p) = \frac{\sum b^x D'_{b_k}}{D'} \text{ and } B^y(p) = \frac{\sum b^y D'_b}{D'}.$$

Consider now a demand curve $D^y_{b_k}(p) = 1 - F_{v|b_k}(p + b^y_k|b_k)$ corresponding to a partial debiasing; namely, elimination of the bias component b^x. Set $D^y = \sum_k D^y_{b_k}$. The reasoning of section 5.3 goes through almost verbatim to establish that

$$B^y(p) \approx \frac{D^Y(p) - D(p)}{D'(p)}.$$

and

$$B^y(p) \approx \frac{D^Y(p) - D(p)}{(D^Y)'(p)}.$$

As shown in the paper, the optimal subsidy satisfies $s^* = B^x + B^y$. It thus follows that if $B^x > 0$—as it would if b_x corresponds to present-biased undervaluations of the energy costs–then $\frac{D^y(p)-D(p)}{(D^y)'(p)}$ constitutes an approximate lower bound for the optimal subsidy. Similarly, if $B^x < 0$—as it would be if b^x corresponds to undervaluations of positive attributes of the incandescent that are not debiased by our interventions—then $\frac{D^y(p)-D(p)}{(D^y)'(p)}$ constitutes an approximate upper bound for the optimal subsidy.