University of Pennsylvania Carey Law School

ILE

INSTITUTE FOR LAW AND ECONOMICS

A Joint Research Center of the Law School, the Wharton School,
and the Department of Economics in the School of Arts and Sciences
at the University of Pennsylvania

RESEARCH PAPER NO. 22-36

When Should Governments Invest More in Nudging? Revisiting Benartzi et al. (2017)

Avishalom Tor
NOTRE DAME LAW SCHOOL

Jonathan Klick
UNIVERSITY OF PENNSYLVANIA CAREY LAW SCHOOL

When Should Governments Invest More in Nudging? Revisiting Benartzi et al. (2017)

Avishalom Tor* and Jonathan Klick**

ABSTRACT

Highly influential recent work by Benartzi et al. (2017) argues—based on comparisons of the respective effectiveness and costs of behavioral interventions (or nudges) versus traditional instruments—that nudges offer more cost-effective means than traditional interventions for changing individual behavior to achieve desirable policy goals. These authors further argue that nudges' cost-effectiveness advantage means that governments and other organizations should increase their investments in such instruments to supplement traditional interventions. Yet a closer look at Benartzi et al.'s (2017) own data and analysis reveals that they variously exclude and include key cost elements to the benefit of behavioral instruments over traditional ones and overstate the utility of cost-effectiveness analysis for policy selection. Once these methodological shortcomings are corrected, a reassessment of key policies evaluated by the authors reveals that nudges do not consistently outperform traditional interventions, neither under cost-effectiveness analysis nor under the methodologically required cost-benefit analysis. These illustrative findings demonstrate that governments should strive to conduct cost-benefit analyses of competing interventions, including nudges, to implement the most efficient of the available instruments.

1. Introduction: The Significance of Benartzi et al. (2017)

Thanks to its prominent authors, venue, and timeliness, Benartzi et al.'s (2017) "Should Governments Invest More in Nudging?" ("the Article") significantly influenced the literature on behavioral regulation in the few years since its publication. The Article was published by a group of leading behavioral economists and behavioral decision researchers (collectively, "the Authors"). These Authors consider whether governments should direct more of their limited budgets to nudging—that is, to using behavioral instruments to achieve their policy goals—a matter of great scholarly and practical interest given the increasing worldwide popularity of this

* **Corresponding Author: Avishalom Tor**, Notre Dame Research Program on Law Market Behavior (ND LAMB), Notre Dame Law School, Notre Dame, Indiana, USA 46556; Israel Institute for Advanced Studies, Safra Campus, Givat Ram, Jerusalem, Israel.

** University of Pennsylvania Carey Law School.

This project benefited from the generous support of Notre Dame Law School and the Israel Institute for Advanced Studies to the first author and the insightful comments of participants at the Hebrew University of Jerusalem Law and Economics Workshop, the 2022 ND LAMB-Tel Aviv University Faculty of Law conference on Empirical, Behavioral, and Experimental Analyses of Law, and the 32nd Annual Meeting of the American Law and Economics Association. Isabella Wilcox and Margarete Tompkins provided excellent research assistance.

approach(Behavioural Insights Team, 2019; EC, 2016; OECD, 2017). It comes as no surprise that since its 2017 publication in *Psychological Science* the Article has been cited more than 700 times by researchers in psychology, economics, public policy, environmental science, health and medicine, energy, technology, law and more.[1]

The Article advances two related arguments: First, it makes the empirical claim that nudging is often more cost-effective than traditional interventions aiming at the same goals. The Authors examine leading studies in four key policy areas, including personal finance, education, energy, and healthcare. In each area, they identify an outcome variable of interest, such as retirement savings or energy conservation and then use cost-effectiveness analysis (CEA) to compare the empirical outcomes high-quality studies obtained for nudges versus traditional interventions targeting that policy outcome. In all four areas, the Article finds that behavioral instruments were more cost-effective than traditional interventions aiming at the same goals. Second, based on this empirical claim regarding the superior cost-effectiveness of nudging, the Authors contend that policy makers should direct more of their limited budgets to behavioral interventions "to supplement traditional policies both inside and outside of governments" (Benartzi et al., 2017: 1052).

As compelling as the Article's claims appear, a closer inspection shows that they are based on methodological oversights that render the Authors' conclusions premature at best. Empirically, Benartzi et al.'s (2017) CEAs exclude key cost categories that ought to be included—most notably the private costs of regulation—even while including as costs resource transfers whose exclusion is required because they do not constitute economic costs (Boardman et al., 2018; Levin and McEwan, 2001). The combined effect of these methodological shortcomings is to make behavioral instruments appear more cost-effective and traditional interventions less cost-effective than the data actually show. When correctly employed, however, the CEA method favored by the Authors reveals no consistent advantage of nudging in the studies they examine.

No less significantly, CEA—which cannot speak to questions of efficiency—is incapable of answering the Article's main normative question (OECD, 2018). To determine whether policymakers should invest more in nudging on an efficiency basis, cost-benefit analysis (CBA) is required (Boardman et al., 2018). While the studies examined by Benartzi et al. (2017) do not offer sufficient information for a full-fledged CBA, together with other public data their findings suffice for illustrative cost-benefit analyses of the energy conservation policies studied by the Authors. These analyses are instructive despite their limitations, revealing that the net benefits of the assessed nudges are, at best, roughly comparable to those of the traditional interventions examined in the Article. The finding that energy conservation nudges are not more efficient than their traditional counterparts leaves the Authors' key normative claim unsubstantiated in an important policy area in which cost-benefit data are readily available.

The substantial impact of the Article renders the softening of its conclusions all the more urgent. By now, the Authors' assertions of the superior cost-effectiveness of nudges and their concomitant advantage as regulatory instruments are routinely accepted as empirically established facts by scholars from across the social sciences (Brandon et al., 2019; Hershfield et al., 2018; Tannenbaum et al., 2017), including even sophisticated behavioral economists (DellaVigna and Linos, 2020). So much so, in fact, that even researchers who criticize nudges on other grounds are mistakenly quick to concede that they "impose nearly zero costs on consumers" (Hagmann et al.,

[1] *E.g.,* Google Scholar counts 704 citations of the Article as of 7/24/2022.

2019), with some commentators going beyond the Authors' own claims erroneously to assert that the Article demonstrated that nudges are sometimes more efficient—not merely more cost-effective—than traditional regulation (De Jonge et al., 2018).

To highlight the Article's methodological oversights and demonstrate how their correction leads to different empirical and normative conclusions from those advanced by the Authors, we focus on the two key policy areas of retirement savings and energy conservation. Part 2 conducts methodologically-appropriate illustrative CEAs of the studies assessed by Benartzi et al. (2017), demonstrating that the evidence in these studies does not show that nudges are more cost effective than traditional interventions. Part 3 then performs an illustrative CBA, finding that the energy conservation nudges examined in the Article do not possess a consistent efficiency advantage over their traditional counterparts. Part 4 concludes, discussing the policy implications of our corrected analyses.[2]

2. Correcting Benartzi et al.'s (2017) CEAs

We revisit the CEAs that underlie the Article's main empirical claim in the two key policy areas of retirement savings and energy conservation interventions. Unlike the Authors' approach, however, our analysis follows the standard methodology of including the private costs of regulation and excluding resource transfers that do not constitute economic costs. Using the same empirical studies upon which the Article relies, the results of this exercise demonstrate that behavioral interventions do not enjoy the dramatic cost-effectiveness advantage over traditional measures asserted by the Authors in either of the assessed policy domains.

2.1. CEA Basics

Cost-effectiveness analysis is a method for identifying the least costly means, per unit of policy benefits, for advancing a given policy goal among a number of available interventions (Levin and McEwan, 2001). This assessment method is widely used in areas such as health and medicine (Miller et al., 2006), education (Levin and Belfield, 2015), energy and the environment (Arimura et al., 2012), and beyond (Boardman et al., 2018).

CEA is often employed when analysts are unwilling to monetize policy benefits, as when these benefits concern the number of lives saved (Layard and Glaister, 1994), or when a policy's impact involves an intermediate good whose monetary value is difficult to assess reliably (Hitch and McKean, 1960), such as the standardized reading test scores of elementary school students. CEA is also employed by regulators tasked with promoting a particular policy goal, who strive to achieve their mandated regulatory targets within limited budgets irrespective of whether their programs are economically justified (Posner, 2003). Allcott (2011: 1088) notes, for example, how "energy conservation program administrators . . . have a set of available programs. In many settings, the administrator will have a regulatory energy conservation target such as an Energy Efficiency Resource Standard that it must achieve using a fixed budget."

As its name indicates, , CEA measures the benefits of competing interventions in terms of policy impact or "effectiveness" to determine their relative appeal. For instance, rather than calculate in monetary terms the private and public benefits generated by increasing college enrollment—the second of the four policy areas examined by Benartzi et al. (2017)—CEA

[2] For the sake of transparency, we report all our calculations in the Online Appendix.

measures how many additional students enrolled in college due to a given policy. Policy costs are then divided by the policy's effectiveness, producing a cost-effectiveness (CE) ratio that allows analysts to compare competing policies. Under CEA, a lower CE ratio—namely, a lower policy cost per unit of effectiveness (e.g., $X per additional college enrollee)—indicates a policy is more attractive than an alternative with a higher CE ratio (Levin & McEwan, 2001).[3]

Even the earliest CEA proponents recognized that analysts must take special care to include all relevant policy costs rather than merely the government's direct budgetary cost to implement a policy (Hitch and McKean, 1960). Indeed, the most significant costs of many policies are their private costs, primarily the opportunity costs generated when interventions change private behavior and thereby cause consumers and firms to forgo whatever benefits they previously enjoyed from their former course of action.[4] Nonetheless, Benartzi et al. (2017) exclude from their CEA all private costs, focusing solely on the implementation costs of the competing interventions. This exclusion makes all of the examined policies appear more cost-effective than they truly are and, importantly, systematically biases the Authors' conclusions in favor of nudges, which typically entail much lower implementation costs than traditional instruments (Tor, 2022, 2023).

In addition, the Article includes in its cost calculations resource transfers that do not constitute economic costs (Posner, 2003). As a result of this inappropriate inclusion, the traditional interventions assessed by the Authors—which primarily rely on financial incentives—seem less cost-effective than they actually are. At the same time, the behavioral instruments to which these traditional interventions are compared involve few resource transfers, so their CEAs are largely unaffected by the inappropriate inclusion of such costs. Once again, therefore, Benartzi et al.'s (2017) approach tilts the cost-effectiveness scales in favor of nudges over traditional instruments.

To illustrate the practical consequences of these oversights, the following sections conduct methodologically appropriate CEAs of the studies on which the Article bases its empirical claim regarding the superior cost-effectiveness of nudging in the two main policy areas of energy conservation and retirement savings. We find that accounting for private costs substantially diminishes the dramatic advantage of nudges claimed by the Authors, while the required exclusion of resource transfers renders transfer-based instruments and behavioral policies roughly comparable in cost-effectiveness terms.

2.2. A corrected CEA of energy conservation policies

Benartzi et al.'s (2017) CEA of energy conservation policies compares two sets of behavioral instruments to two studies of traditional, primarily financial interventions. On the behavioral side, the Authors include Allcott's (2011) study of a social information nudge and Asensio and Delmas's (2015) social, environmental and health information nudge, which they compare to a conditional rebate policy examined by Ito (2015) and Arimura et al.'s (2012) large-scale analysis of hundreds of demand-side management (DSM) programs that largely relied on

[3] Less commonly, analysts sometimes calculate the reciprocal effectiveness-cost (EC) ratio instead, as in Benartzi et al. (2017). Although identical in terms of their conclusions, EC ratios describe policies in terms of their effectiveness per unit of cost (e.g., the number of additional college enrollees per $1000). When using the reciprocal EC ratio, a policy is more attractive the higher its effectiveness per unit of cost. For consistency with the broader literature, however, the analysis here uses the standard CE (rather than EC) ratio, converting the Article's figures as necessary.

[4] Naturally, there are also opportunity costs on the government side of the ledger when it implements one policy over its competitors—namely, the forgone welfare benefits of the government's best alternative use of the same budget towards another intervention (Pearce, 1983; Sugden and Williams, 1978).

financial incentives. The Authors find that the behavioral instrument of Allcott (2011)—the far more effective of the two nudges—is about twice as cost-effective as the traditional programs studied by Arimura et al. (2012) and seven times more so than Ito's (2015) policy. Thus, as reflected in *Panel A* of Table 1, the CE advantage of the more cost-effective behavioral instrument appears substantial. However, a closer look at the cost figures used by Benartzi et al. (2017) reveals that the dramatic advantage of Allcott's (2011) nudge is due to the Authors' inappropriate exclusion of private costs and inclusion of resource transfers in their cost calculations (Boardman et al., 2018; Levin and McEwan, 2001). In what probably was the first large-scale randomized controlled trial (RCT) to test the effectiveness of an energy conservation nudge, Allcott (2011) studied the effects of Home Energy Reports (HERs)—which provide targeted households with social comparisons of their energy use to that of their neighbors, suggest social norms favoring energy conservation, and offer simple energy-saving tips—on the electrical consumption of about 600,000 residential homes, finding an average treatment effect ranging from 1.4% to 3.3%, with an unweighted mean of 2.03%. Based on these findings, Allcott (2011) concluded that, when dividing the administrative implementation cost of printing and sending the reports to the targeted households by the average kWh saved per year, the programs studied had an unweighted mean CE of $0.04/kWh saved.

The Article also considers a smaller-scale study by Asensio and Delmas (2015), which examined the effects of two nudges that provided recipients with social information on their energy consumption while quantifying its relative private costs or negative environmental effects (with a general linking to negative health effects). This intervention combined a specialized website and a sophisticated wireless sensor network that monitored electrical usage and provided recipients with detailed energy consumption information for major appliances. In addition, treated households received weekly information that compared their costs or emissions to the most efficient among their neighbors, with emission reports also highlighting their negative health effects.

Benartzi et al. (2017: 1049) report, based on personal communication, that Asensio and Delmas's (2015) nudges entailed a very high implementation cost of $3,532.23 per household, largely due to the substantial costs of installing the sophisticated sensor network and developing the specialized website (Asensio and Delmas, 2015). The effectiveness of the treatments was mixed, with the private costs nudge producing a non-significant increase and the environmental and health nudge yielding a substantial 8.22% decrease in energy use that translated to a 71.2 kWh consumption reduction over the 100-day trial. To render these results more comparable to the annual energy savings figures reported for competing interventions, the Authors assumed that the impact of the nudge decayed linearly over the remainder of one year, arriving at an annual effectiveness figure of 149.8 kWh per household. Consequently, the Article concludes that the cost-effectiveness of the latter nudge was a very unappealing $23.58/kWh saved.

Benartzi et al. (2017) favorably compared Allcott's (2011) results to the cost-effectiveness of the traditional electricity conservation interventions studied by Ito (2015) and Arimura et al. (2012). Ito (2015) examined one of the largest electricity subsidy (rebate) programs in the United States in terms of expenditure and the number of customers who received rebates, which California implemented during the summer of 2005. His study found that the program produced an average reduction of 17.34 kWh per household overall, at an implementation cost of $5.96 per customer ($4.33 for rebates plus $1.63 for administrative and marketing costs), indicating this policy had an overall CE of $0.34/kWh saved.

Notably, however, Ito's (2015) overall \$0.34/kWh CE figure—nearly ten times the CE reported for Allcott's (2011) nudge—averages the dramatically different levels of effectiveness of the rebate program over two groups of California electrical power customers. The study found a negligible treatment effect of 3.11 kWh for coastal customers even while the program achieved a significant effect of 169.15 kWh in inland areas. Naturally, these disparate effects resulted in two very different CE ratios: A costly \$1.86/kWh saved in coastal areas, but an inland CE of \$0.04/kWh saved comparable to the CE of Allcott's (2011) nudge.

The comprehensive study by Arimura et al. (2012) examined the long-term cost-effectiveness of widely-used electricity conservation interventions that the energy literature refers to as “demand side management” (DSM) programs, which rely on financial incentives and energy conservation tips. The authors analyzed a panel of annual utility-level DSM data from 1992 to 2006 across the U.S. that included a main sample of 3,326 utility-level observations from 307 utilities. Following Arimura et al. (2012), the Article pegs the cost-effectiveness of these programs at \$0.08/kWh saved when accounting for the longer-term benefits of these interventions (e.g., through encouraging household investments in home improvements) and assuming a 5% discount rate (implying annual energy savings of 152.54/KWh per customer given an implementation cost of \$12.67 per customer). This estimate is twice as high as Allcott’s (2011) CE figures, thereby appearing to support Benartzi et al.’s (2017) conclusion that energy saving nudges can be more cost effective than traditional DSM measures.

Yet, as noted earlier, Benartzi et al.’s (2017) calculations ignore the private costs of energy conservation policies, even while including resource transfers that do not constitute economic costs. *Panel B* of Table 1 reports an illustrative CEA that corrects for these methodological oversights, showing that in fact the most cost-effective energy conservation nudge does not outperform traditional DSM programs in the studies the Article evaluates.

[PLACE TABLE 1 ABOUT HERE]

2.2.1 Private costs: Consumer costs and retailer net revenue loss

Since Benartzi et al. (2017) do not assess the private costs of energy conservation policies, we draw on the findings reported by the studies assessed in the Article, together with more recent research, to offer conservative estimates of the likely range of these private costs. These estimates illustrate how a methodologically-appropriate CEA does not support the Article's main empirical claim regarding the CE superiority of nudges over traditional DSM policies.

We are not the first to recognize the limits of approaches that ignore the private costs of energy conservation nudges (Tor, 2023). Indeed, Allcott (2011) already acknowledged that his cost-effectiveness calculations exclude consumer costs (CCs)—an important category of private costs—and thereby understate the true cost of the HERs. To address this and further questions regarding the welfare effects of HERs, Allcott and Kessler (2019) more recently conducted a CBA of a program that sent four HERs to natural gas consumers in the state of New York over the 2014-2015 heating season. They estimated the average non-energy costs of the nudge to recipient households at \$2.46 or 43% of the households' savings from reducing their natural gas consumption.

These private costs likely included both the direct costs and the opportunity costs of the nudge. The direct "moral cost" of the HERs is the utility loss to consumers from repeated nudging (e.g., due to being subjected to upwards social comparisons, to the implication that they violate injunctive energy conservation norms, or perhaps even to mere annoyance), while their opportunity costs are the forgone benefits from their formerly higher energy consumption (such as the enjoyment from a more comfortable indoor temperature in the winter and the summer or from not needing to remember to turn off lights in temporarily unused rooms) (Tor, 2020).

The distinction between these two potential sources of CCs is important for estimating the cost-effectiveness of the electricity HERs in Allcott (2011). The direct costs of the reports are likely to be roughly similar irrespective of the source of energy consumption they target (i.e. natural gas versus electricity). On the other hand, the HERs' opportunity costs, which reflect the benefits that consumers sacrifice to reduce their energy use, should be more proportionate to the benefits of energy conservation—that is, the greater the conservation benefits to consumers, the higher the opportunity costs they will bear willingly to obtain these benefits.

Hence, if the non-energy costs to recipient households of the natural gas nudge were primarily due to its direct costs, the $2.46 per household of Allcott and Kessler (2019) may constitute a reasonable if low estimate of the magnitude of the CCs of Allcott's (2011) electricity HERs.[5] However, if the HERs' main non-energy costs to recipient households are their opportunity costs, they could amount to $14.60 (43% of average retail savings of $33.95) for Alcott's (2011) electricity nudge.[6] The fourth column of *Panel B* in Table 1 reports these and our other CC estimate for the competing energy conservation policies.

In a similar vein, Asensio and Delmas's (2015) design did not allow for direct estimates of the CCs to participating households.[7] Nonetheless, since this intervention was essentially a more heavy-handed variant of the standard HER in both messaging frequency and moral tone, it was unlikely to have generated lower private costs than those produced by standard HERs. We therefore apply here the same approach as for Allcott's (2011) nudge, which likely still underestimates the present CCs, arriving at $2.46 or $10.31 per customer (43% of average consumer electricity retail savings of $23.97).

Turning to the studies of traditional interventions examined by Benartzi et al. (2017) in this area, neither Ito (2015) nor Arimura et al. (2012) provide CC estimates, though they do note their potential significance for policy assessment. Thus, for illustrative purposes, we follow the benchmark estimate cited by both Allcott and Greenstone (2012) and Arimura et al. (2012), which sets the CCs of traditional DSM interventions at approximately 70% of program implementation

[5] This would be a conservative estimate even under the assumption that the nudge only entailed direct consumer costs because Allcott's (2011) electricity nudge on average involved sending about eight reports on average over a whole year, compared to sending only four reports during one heating season (of 243 days) in Allcott and Kessler's (2019) natural gas nudge.

[6] A similar approach is also suggested by Allcott and Kessler's (2019) "speculative" assessment of a typical electricity HER in their Online Appendix. Notably, moreover, even this higher figure is a conservative estimate of the true CCs of the nudge, since Allcott and Kessler (2019) also report that consumers may have dramatically overestimated their savings from energy conservation. If this were the case, consumers' willingness to pay for the HERs would have been biased upwards and the reports' actual consumer costs would have been substantially higher than the present estimate of 43% of energy savings.

[7] Asensio and Delmas's (2015) findings indicate that these costs were substantial, since participants exhibited a non-significant *increase*—rather than any decrease—in energy consumption when faced with a nudge that provided detailed and frequent information about their private electricity costs, despite having expressed their wish to conserve energy in a pre-treatment survey.

costs.[8] Using this benchmark, CCs for the California rebate program assessed by Ito (2015) are $4.17 overall, $4.07 for coastal customers, and $5.32 for inland customers, while for Arimura et al.'s (2012) comprehensive longitudinal study of DSM programs CCs are $8.87.

In addition to their consumer costs, however, policies that successfully reduce residential electricity usage also entail commensurate revenue losses for electricity retailers. A substantial portion of this revenue loss represents the retailers' avoidable costs—most notably, their own energy acquisition costs[9]—which do not constitute economic costs. But the remaining cost of reduced retail electricity consumption represents retailers' net revenue loss ("RNRL"), a private cost that CEAs of competing interventions in this area should account for.

Although Allcott (2011) does not report RNRL data, Allcott and Kessler (2019) estimate in their Online Appendix the RNRL for a typical long-term electricity HER program at 25% of retail electricity cost savings. This figure is substantially lower than Baskette et al.’s (2006)modal estimate of avoidable operating costs at 42% of retail prices across all California DSM programs in 2004, which implies a modal RNRL at 58% electricity retail prices—reported by. Hence, for the nationwide studies of Allcott (2011) and Arimura et al. (2012) we retain Allcott and Kessler's (2019) typical RNRL estimate of 25% of electricity retail prices, while applying Baskette et al.'s (2006) California figures to the studies of Asensio and Delmas (2015) and Ito (2015), both of which took place in California.[10]

The fifth column of *Panel B* in Table 1 reports these estimates, finding that the nudges of Allcott (2011) and Asensio and Delmas (2015) entailed RNRL of $8.49 (at 25% of retail electricity cost savings) and $13.90 (at 58% of retail electricity cost savings), respectively. The RNRL of the California rebate examined by Ito (2015) was $1.91 overall, $0.34 for coastal customers, and $18.64 inland customers (at 58% of retail electricity cost savings), while that obtained for the DSM programs in Arimura et al. (2012) was $3.17 (at 25% of retail electricity cost savings).

2.2.2 Total costs

Using our private costs estimates, we can calculate the total costs of the energy conservation policies examined by Benartzi et al. (2017) as Total Costs = Implementation Costs + Private Costs, with Private Costs = Consumer Costs + RNRL. This calculation yields total policy costs per customer of $20.67 or $32.81 for Allcott's (2011) nudge and $3,548.59 or $3,556.44 for Asensio and Delmas's (2015) nudge. Total costs for Ito's (2015) rebate program are $12.04 overall, $10.22 for coastal customers, and $31.56 for inland customers, while Arimura et al.’s (2012) total costs are $24.71.

2.2.3 Corrected cost-effectiveness I (including private costs)

We can now assess the cost-effectiveness of the competing energy conservation policies in the Article by dividing Total Policy Cost per customer by Effectiveness (average usage reduction per customer). The results, reported in the sixth column of *Panel B* in Table 1, show Allcott’s (2011) nudge with a corrected CE ratio of $0.09 or $0.14/kWh saved, and the nudge of Asensio

[8] While Arimura et al. (2012) only cites this benchmark to note how CE calculations that ignore private costs understate actual DSM costs, Allcott and Greenstone (2012) employ it to assess the overall cost-effectiveness of DSM programs.
[9] Retailers' avoidable costs include additional components, such as transmission and distribution capacity costs (e.g., Baskette et al., 2006), but these additional costs tend to make only a small fraction of avoidable costs.
[10] Applying the typical non-California RNRL rate to the California programs does not appreciably change the conclusions of our CE comparisons.

and Delmas (2015) with a corrected CE ratio of $23.69 or $23.74/kWh saved. The traditional policy of Ito (2015) has a CE ratio of $0.69/kWh saved overall, $3.29/kWh saved for coastal customers, and $0.19/kWh saved for inland customers, while the DSM programs studied by Arimura et al. (2012) had a CE ratio of $0.16/kWh saved.

A corrected CEA that accounts for the total costs—both public and private—of the energy conservation policies examined by Benartzi et al. (2017) therefore suggests that the better-performing nudge of Allcott (2011) remains far more cost-effective than the conditional rebate in Ito (2015) overall, but that performance gap is much narrower when compared to Ito's (2015) inland customers. More significantly, though Allcott's (2011) nudge retains an advantage over the DSM programs in Arimura et al. (2012) under the lower estimate of the HERs' consumer costs, this advantage diminishes under the higher (and likely more realistic) CCs estimate.

2.2.4 Corrected cost-effectiveness II (also excluding transfers)

While the corrected CE figures that account for both the public and the private costs of energy conservation policies illustrate the importance of accounting for private costs—which Benartzi et al. (2017) overlook—a methodologically-appropriate CEA must also exclude resource transfers, which do not constitute true economic costs. This further-corrected CE demonstrates that the most cost-effective energy conservation nudge considered by the Article is not much more, and possibly even less, cost-effective than traditional DSM programs.

The seventh column of *Panel B* in Table 1 shows that the exclusion of resource transfers does not affect the cost figures for the nudges of Allcott (2011) and Asensio and Delmas (2015), since neither entailed such a transfer. The preceding CE figures for these nudges remain unchanged, with CE ratios of $0.09 or $0.14 per kWh saved for the former and $23.69 or $23.74 per kWh saved for the latter.

This is not the case, however, for the traditional energy conservation interventions, whose implementation costs were primarily transfers from electricity suppliers' general customer base (in the form of higher electricity rates) to the households that received the incentives offered by DSM programs. Estimates of the proportion of program implementation costs that constitutes economic costs—such as their administrative, marketing, or delivery costs—vary greatly among DSM programs. While Wirtshafter Associates (2006) estimated that these costs were as low as 13% of California's 2005 rebate programs' overall implementation costs in its report to California's three large utility companies, Ito (2015) estimated the same costs at 27% of implementation costs for the California rebate program that he examined, and recent estimates suggest that such costs may even amount to as much as 39% of the implementation costs of all DSM programs in 2020 (Minor, 2019). For consistency, we apply the 27% estimate of Ito (2015)—which is also about midrange between the lower and higher estimates of the other studies noted above—to both Ito's (2015) own data and Arimura et al.'s (2012) DSM findings. This estimate produces non-transfer implementation costs of $1.63 for Ito (2015) and $3.42 for Arimura et al. (2012).

The eighth column of *Panel B* in Table 1 reports the results of a further-corrected CEA that also excludes the portions of the competing policies' implementation costs that were mere transfers. As already noted, the CE ratio of the two nudges remains unchanged, while the further-corrected total cost figures of Ito's (2015) rebate program amount to $7.71 per customer overall, $6.04 per customer for coastal households, and $25.59 for inland households, which yield CE ratios of $0.44/kWh saved overall, $1.94/kWh saved for coastal households, and $0.15/kWh saved for inland households. A similar further correction of Arimura et al.'s (2012) DSM figures shows these programs entailed total non-transfer costs of $15.46 per customer, generating a further-corrected CE ratio of $0.10/kWh saved.

Therefore, a further-corrected, methodologically-appropriate, CEA that both includes private costs and excludes transfers shows Allcott's (2011) HER nudge to be within the cost-effectiveness range of the traditional DSM programs studied by Arimura et al. (2012), though slightly more or notably less cost-effective than these common instruments, depending on the CC estimate selected for the nudge. Asensio and Delmas's (2015) nudge still performs poorly in terms of cost-effectiveness, due to its exceedingly high implementation costs. Because of its very low effectiveness among coastal households, the California rebate program studied by Ito (2015) also continues to underperform both other traditional instruments and the more successful nudge, although the program's performance among inland households is resembles the CE ratio of DSM policies more generally.

2.3. A corrected CEA of retirement savings policies

The corrected assessment of retirement savings interventions presents somewhat different issues from those examined above, since the very goal of such policies is to increase the transfer of resources into retirement savings accounts. This is the case when non-consumer transfers—such as tax subsidies, matching contributions, or other incentive schemes—are used to increase individuals' retirement savings. But it is also the case whenever people contribute to their retirement accounts from their own resources that they inevitably transfer from current consumption or other savings.

According to the standard approach, therefore, all interventions that increase retirement savings transfers will appear inefficient under CBA. They produce no economic benefits, but entail certain (non-transfer) implementation costs as well as some direct private costs for their targets, who must make difficult decisions concerning their retirement savings (Goldin et al., 2020) or expend resources to execute these decisions, thereby rendering standard cost-benefit evaluations inapposite. Nevertheless, comparisons of the cost-effectiveness performance of competing retirement savings policies can be somewhat informative insofar as they illuminate their relative implementation costs per unit of effectiveness (Posner, 2003).

Benartzi et al.'s (2017) CEA of competing retirement savings policies compares a nudge studied by Carroll et al. (2009) to four traditional interventions, three of which the Article characterizes as financial-incentive instruments (Chetty et al., 2014; Duflo et al., 2006; Duflo et al., 2007) and one as an educational policy (Duflo and Saez, 2003). The Authors find a dramatic cost-effectiveness advantage of the nudge over its traditional competitors, with the former exhibiting a CE of merely $0.01 per $1 of increased contributions and the latter showing CE ratios ranging from $0.07 (Duflo and Saez, 2003) to $0.81 (Duflo et al., 2007) per $1 of increased retirement savings contributions. In other words, according to the Article's calculations, the Caroll et al. (2009) nudge was seven times more cost effective than the best-performing traditional intervention (Duflo and Saez, 2003), and almost two orders of magnitude more cost-effective than the worst-performing among them (Duflo et al., 2007). On its face, therefore, this analysis indicates that nudging possesses an overwhelming CE advantage as a means for encouraging employees to save more for retirement.

Once again, however, a Benartzi et al.'s (2017) cost figures inappropriately exclude private costs and include public resource transfers. The first oversight is less detrimental in the present context, in which intrapersonal transfers (rather than true economic costs) constitute the bulk of private costs. The competing retirement savings interventions in the Article do entail some non-transfer private costs (Goldin et al., 2020; Tor, 2022), but these costs cannot be estimated reliably with the available data and are unlikely to change our qualitative conclusions.

Yet this is not the case with Benartzi et al.'s (2017) inclusion of non-consumer resource transfers in their implementation cost figures. This error systematically tilts the Article's cost estimates in favor of the assessed nudge over the financial instruments targeting retirement savings, since the former entailed only some administrative implementation costs and involved no non-consumer transfers, , while the main implementation costs of latter were non-consumer transfers. A CEA that is corrected to exclude these transfers, as methodologically required, therefore produces very different CE ratios from those calculated by the Authors.

Before turning to the results of our corrected CEA, we should note that the Article consciously disregards the non-transfer implementation costs of the traditional policies it assesses. Since Benartzi et al. (2017: 1044) include transfers in their cost calculations, they decided that any remaining "minor unreported" costs can be overlooked, explaining that "for grant programs or tax credits, administrative and marketing costs are small in relation to the total amount of money transferred, so accounting for them would not significantly affect…estimates" (1053). Consequently, as the Authors acknowledge, they ended up accounting for non-transfer implementation costs for the retirement savings nudge—where they did not deem them "minor"—but ignored them when assessing the competing traditional instruments.

Once resource transfers are appropriately excluded, however, those "minor unreported" costs are the only remaining implementation costs that constitute true economic costs. To illustrate how the results of a corrected CEA differ from those reported in the Article, we must therefore estimate the administrative, marketing, or similar non-transfer costs of the competing retirement savings policies. In doing so, we follow Benartzi et al. (2017), who made a similar rough estimate of the administrative costs of Carroll et al.'s (2009) nudge. Moreover, though our illustrative estimates suffice to demonstrate the significance of non-transfer costs for the comparative assessment of retirement savings policies, our qualitative conclusions are robust to alternative reasonable estimates of these costs.

Table 2 reports the results of our illustrative corrected CEA, which divides the non-transfer implementation costs of each retirement savings policy by its effectiveness. The nudge studied by Carroll et al. (2009) concerned a "required choice" policy that asked employees, upon hiring, to choose their preferred contribution rate within 30 days, but left them free to decide whether and how much to contribute and imposed no penalty on those who failed to make the choice. The Authors estimate the costs of the policy to be $2.34 of administrative expenses per employee, for adding a form to the company's hiring packet and making follow-up phone calls to those who did not make the choice despite receiving the form. The nudge yielded a $234 average increase in annual employee contributions, so Benartzi et al. (2017) concluded that its CE was $0.01/$1 of increased contributions. Furthermore, because the nudge did not entail a resource transfer, its CE ratio remains unchanged under our corrected CEA.[11]

[11] We retain the Authors' estimate of the effectiveness of the Carroll et al. (2009) nudge, which was conservative—treating the 1% increase in contributions it found (in a sample whose median income was approximately $30,000 in 1998 dollars) as if it applied to lower-income employees with a median income of $20,000 in 2015 dollars—and consequently may have understated the policy's cost-effectiveness ratio.

[PLACE TABLE 2 ABOUT HERE]

With respect to the traditional interventions examined in the Article, however, the case is quite different. The elimination of the Danish tax subsidy to one type of retirement savings account in Chetty et al. (2014) reduced the affected individuals' contributions by $632, while raising government revenues by $228.[12] This led the Authors to find that the policy change produced an uncorrected CE ratio of $0.36/$1 of change in retirement savings contributions. But a corrected CEA must exclude the transfer of $228 in taxes from the affected taxpayers to the Danish government while including the non-transfer implementation costs to both the government and employers of adjusting to the policy change (Chetty et al., 2014). For illustrative purposes, we estimate these costs at $2.00 per affected taxpayer, with a resulting corrected CE ratio of $0.003/$1 decrease in contributions.[13]

The Article finds an uncorrected CE ratio of $0.07/$1 of change in retirement savings contributions for the education and incentives policy of Duflo and Saez (2003), which increased annual contributions by $68.97, at implementation costs of $4.73 per affected employee. However, these costs should be excluded since they were transfer payments—in the form of small monetary incentives to encourage employees to attend the employer's benefit fair—rather than economic costs. At the same time, we estimate the non-transfer implementation costs of a reminder letter that Duflo and Saez (2003: 823) also produced and sent a to half of the participants at $2 per letter or $1 per employee, with a resulting corrected CE of $0.01/$1.00 retirement savings increase.

Duflo et al. (2006), the third of the traditional interventions Benartzi et al. (2017) compare to Carroll et al.'s (2009) nudge, involved a program that offered matching incentives of 20% or 50% of IRA contributions to low- and middle-income families. The effectiveness of the 20% match was $109.51 of increased annual contributions at an implementation cost of $19.54 per affected taxpayer, while that of the 50% match was $286.07 of the same at an implementation cost of $96.41 per affected taxpayer, with resulting uncorrected CE ratios of $0.18 and $0.34/$1.00 increase in retirement savings contributions for the 20% and 50% matching levels respectively. Here too, the matching expenditures were mere transfers—in this case from a tax-preparation company to taxpayers—that should be excluded from our cost figures.

The matching program also entailed substantial non-transfer implementation costs—most notably the specialized training sessions, materials, and tax-preparation software modifications that the company provided to its tax professionals, but also the additional time these professionals spent with relevant clients to explain and discuss the matching opportunity (Duflo et al., 2006: 1316-1317), which the Article ignored. For illustrative purposes, we estimate these costs at $5.00

[12] For present purposes, we follow Benartzi et al.'s (2017) treatment of Chetty et al.'s (2014) findings as if they are comparable to the other retirement savings interventions the Authors assess. However, the real effectiveness of the Danish tax policy change was negligible. Chetty et al. (2014) report that the change, which affected only one of two types of tax-privileged accounts, led affected taxpayers to divert 47% of their affected contributions (1154 DKK) into another tax-privileged vehicle and nearly all of the remaining 53% (1295 DKK), post-tax, into non-privileged savings accounts (518 DKK). Ultimately, only 0.1% of the diverted savings—or 3 DKK—were left as newly disposable income in lieu of retirement savings). Using this figure, the effectiveness of the policy was in fact only $0.75 per affected taxpayer.

[13] In principle, these non-transfer implementation costs should be included in the basic implementation costs figure. For expositional purposes, however, we follow Benartzi et al's (2017) approach of ignoring these costs in the basic estimate, as including these costs does not change the uncorrected CE figures.

per affected individual, which produces corrected CE ratios of \$0.05/\$1.00 and \$0.02/\$1.00 increase in retirement savings contributions for the 20% and 50% matching levels respectively.

Last, the effectiveness of the federal income tax Saver's Credit program targeting low- to middle-income households that Duflo et al. (2007) studied was \$13.56 in increased annual contributions with \$10.85 in implementation costs per affected taxpayer, which yielded an uncorrected CE ratio of \$0.80/\$1.00 increase in retirement savings contributions. Since the federal tax credit was a transfer rather than an economic cost, it should be excluded. At the same time, the program must have entailed some minor non-transfer government implementation costs of adjusting the federal tax assessment process, which we estimate at \$0.50 per affected taxpayer for illustrative purposes. Hence, the Saver's Credit program had a corrected CE ratio of \$0.04/\$1.00 increase in retirement savings contributions.

All in all, our illustrative corrected CEA of the competing policies in this area finds that not only does Carroll et al.'s (2009) nudge not hold the dramatic advantage over its traditional competitors that the Article suggests, but some financial incentive policies perform as well as and likely better than this nudge . Importantly, the CE performance differences among these competing transfer policies are sensitive to the magnitude of their non-transfer costs, which can be low for nudges and traditional financial incentive instruments alike, as well as to the effectiveness of the competing interventions in increasing retirement savings contributions.

3. An Illustrative CBA of Benartzi et al.'s (2017) Energy Conservation Policies

Even a methodologically-appropriate CEA cannot answer the Authors' main normative question—namely, whether governments should invest more in nudging—for which we must turn to cost-benefit analysis instead.

3.1. CBA versus CEA

By considering the costs of different interventions rather than focusing on their effectiveness alone, CEA plays a valuable role in policy assessment, but its utility is limited in two crucial respects: For one, CE comparisons only show which policy provides regulators with the greatest unit "return on investment" rather than which policy is more efficient, so such comparisons can erroneously support the selection of a less efficient policy that is more cost effective. Furthermore, by starting from the (implicit) assumption that one of the available interventions must be implemented, CEA can favor the selection of inefficient interventions (Boardman et al., 2018).

Importantly, this fundamental limitation of CEA goes beyond its inability to "address…whether increasing the behavior in question is socially beneficial," which Benartzi et al. (2017: 1044) correctly recognize, because agreement regarding the social benefits of changing a particular behavior does not imply that any of the currently available means for accomplishing this goal is capable of producing net social benefits. Thus, even the Authors' own approach of "tak[ing] government goals as given and then…address[ing] how best those goals can be achieved" still necessitates a cost-benefit analysis rather than CEA.

CBA is the dominant approach to policy assessment worldwide: It is mandated for U.S. federal regulation by a series of Executive Orders (e.g., Federal Register, 1993) and plays an important role in the mandatory regulatory impact assessment processes of the European Union (EC, 2021), other OECD countries (OECD, 2020) and beyond (De Francesco, 2012; Dunlop and Radaelli, 2016). Because inefficient policies are socially harmful, while taking the most efficient available action maximizes social welfare, CBA—Unlike CEA—strives to identify the most efficient instrument available in a given policy context (Boardman et al., 2018).

To do so, CBA quantifies the social consequences of policy interventions in monetary terms. The basic conceptual framework of CBA is straightforward: From the perspective of efficiency, the value of a policy to society is measured by its net social benefits—that is, the public and private benefits it generates minus its public and private costs (Layard and Glaister, 1994). Based on this assessment, CBA directs policy makers to select the option that offers the highest net social benefits and to avoid inefficient policies that fail to offer any net social benefits vis-à-vis the status quo. The net-benefits metric therefore directs attention beyond considerations of effectiveness or even cost-effectiveness, to the overall economic effects of a policy.[14]

3.2. An Illustrative CBA of energy conservation policies

A CBA of energy conservation policies weighs the social benefits of such interventions against their social costs to determine the magnitude of the net benefits they provide, if any. Importantly, energy conservation entails serveral public and private benefits and costs: Reductions in electricity consumption produce public benefits by reducing harmful externalities and private benefits by lowering household expenditures, but conservation policies entail public implementation costs (excluding transfers)and the private costs of both retailer net revenue losses from diminished electricity sales and consumer costs (both direct and opportunity costs).

Our corrected CEA of policies targeting energy conservation in Part 2.2 already provides estimates for the most of the benefits and costs of the assessed policies, with the exception of externality reduction benefits. The literature offers a variety of estimates of electricity production externalities, such as the global median externality cost of $0.025/kWh finding of Sovacool et al.'s (2021) recent meta-analysis and research synthesis of existing studies, but our illustrative benefit estimate follows the slightly higher $0.033/kWh figure of Allcott and Kessler (2019) and their Online Appendix.[15]

Table 3 shows the net benefits of the policies assessed by Benartzi et al. (2017). *Panel A* reports net welfare other than consumer welfare ("non-consumer welfare"), which equals the benefits of externality reduction minus the non-transfer implementation costs and retailer net revenue loss. Our illustrative estimate finds the externality benefits generated by the reduction in electricity consumption due to the assessed interventions at $7.47 for Allcott's HERs (2011) and $4.94 for Asensio and Delmas's (2015) nudge. For the traditional instruments, Ito's (2015) rebate program generated externality reduction benefits of $0.57 overall, $0.10 for coastal customers, and $5.58 for inland customers, while the externality reduction benefits of Arimura et al.'s (2012) DSM policies were $5.03.

We calculate the net non-consumer welfare effects of the different polices by subtracting their non-transfer implementation costs and the retailer net revenue loss from the externality

---

[14] The literature discussing the merits, demerits, and challenges of cost-benefit analysis is extensive, and outside the scope of the present, focused, reassessment of Benartzi et al. (2017) (e.g., Adler and Posner, 2001, 2006; Boardman et al., 2018; Layard and Glaster, 1994; Pearce, 1983; Sunstein, 2018; Weimer, 2017).

[15] Sovacool et al. (2021) also estimate mean global externalities of electricity production at $0.079/kWh—more than twice as high as our main estimate. This figure, which averages global externalities, including countries like China or India that produce substantially higher mean electricity generation externalities, does not reflect the externality benefits of the domestic U.S. energy saving interventions assessed here. Moreover, using this higher figure would increase the assessed external reduction benefits of all competing policies in proportion to their energy savings.

reduction benefits. This calculation finds that Allcott's (2011) nudge produced net non-consumer welfare of -$10.74, while Asensio and Delmas's (2015) costly nudge produced net non-consumer welfare of -$3,541.19. The California rebate program studied by Ito (2015) produced net non-consumer welfare of -$2.97 overall, -$1.87 for coastal households, and -$14.69 for inland households. Arimura et al.'s (2012) DSM programs produced net non-consumer welfare of -$1.56. Hence, all of the assessed policies generated net non-consumer costs, though the better-performing nudge of Allcott (2011) still produced higher net non-consumer costs than the traditional financial-incentive policies due to the higher non-transfer implementation costs it entailed.[16]

[PLACE TABLE 3 ABOUT HERE]

Panel B of Table 3 reports the net consumer welfare effects of the assessed policies, which we calculate by subtracting the consumer costs of reduced electricity consumption (already reported in *Panel B* of Table 1) from consumers' retail electricity savings. Allcott's (2011) nudge produced net consumer welfare of $31.49 or $19.35, and Asensio and Delmas's (2015) costly nudge produced net consumer welfare of $21.51 or $13.66. Ito's (2015) rebate program produced net consumer welfare of -$0.88 overall, -$3.48 for coastal households, and $26.82 for inland households. Arimura et al.'s (2012) DSM programs produced net consumer welfare effect of $11.50. Thus, under our estimates, most interventions produced positive net consumer welfare, with the exception of Ito's (2015) program overall and inland households.

Most important, the net social welfare effects of the energy conservation policies examined by Benartzi et al. (2017) are reported in the last column of Table 3, *Panel B* . Allcott's (2011) nudge yielded net social welfare of $20.75 or $8.61, while that of Asensio and Delmas (2015) was -$3,519.68 or -$3,524.43. Ito's (2015) rebate program generated a net effect of -$3.85 overall, -$5.35 for coastal households, and $12.13 for inland households; and Arimura et al.'s (2012) programs produced a net social welfare effect of $9.94.

Our illustrative CBA assesses the welfare effects of the energy conservation interventions under consideration, a matter that the CEA employed by Benartzi et al.'s (2017) is incapable of illuminating and without which the Authors key normative query of whether governments should invest more in nudging cannot be answered. This exercise reveals, for instance, that both the Asensio and Delmas (2015) nudge and Ito's (2015) overall rebate program produced net social welfare losses and should not have been adopted. Indeed, even the best-performing among the competing policies generated very modest net social benefits, with the more optimistic estimate of Allcott's (2011) HERs yielding only about $0.40 per customer household per week.

According to our illustrative CBA, moreover, the identity of the most efficient energy conservation intervention—Allcott's (2011) HERs or Arimura et al.'s (2012) DSM policies—depends on one's view regarding the best estimate of the consumer costs of the HERs. Specifically, the less optimistic estimate of these costs puts the net social benefits of Allcott's (2011) nudge slightly below the traditional DSM programs in Arimura et al. (2012), while the more optimistic estimate of the HERs consumer costs finds their net benefits about twice as large as those of the latter.

[16] This result is in line with the recent findings of Allcott and Kessler's (2019) CBA of a natural gas HER nudge and their Online Appendix estimate of the net non-consumer welfare effects of one of the four-year electricity HERs studied by Allcott and Rogers (2014).

For present purposes, we need not decide the matter; the reality of these two plausible conflicting conclusions is itself informative. To wit, in contrast to the seemingly clear superiority of Allcott's (2011) nudge according to Benartzi et al.'s (2017) flawed and inapposite CEA, the present CBA suggests that further investigation into the consumer costs of the nudge is needed to determine more conclusively whether Allcott's (2011) HERs or Arimura et al.'s (2012) traditional DSM policies are the more efficient.

4. Conclusion

The methodologically-appropriate CEAs of the interventions evaluated in the Article demonstrate that Benartzi et al.'s (2017) basic empirical finding—namely, that nudges are far more cost-effective than traditional policies—is overstated. In the energy conservation area, instead of being doubly more cost-effective than the common DSM policies studied by Arimura et al. (2012), Allcott's (2011) best-performing nudge turned out to be roughly as or somewhat less cost-effective. Similarly, rather than being seven to eighty-one times more cost-effective, the retirement savings nudge of Carroll et al. (2009) was shown to be somewhere between five times less and five times more cost-effective than its various traditional competitors.

While these results already paint a significantly different picture from that advanced by the Article, our illustrative CBA of the same energy conservation policies assessed by the Authors offered further observations beyond what even a methodologically-appropriate CEA could provide: For one, it indicated that the implementation of two of the assessed policies—namely, Asensio and Delmas (2015) nudge and Ito's (2015) overall rebate program—was inefficient and thus socially harmful. Second, our CBA showed that either Allcott's (2011) nudge or the traditional instruments in Arimura et al. (2012) could be the most efficient among the assessed interventions.

Taken together, these analyses demonstrate that nudges are neither systematically more cost effective nor consistently able to produce larger net social benefits than traditional regulatory instruments, at least in the major policy areas examined here. This should not be particularly surprising once we consider the nature and full range of costs that policy interventions generate on both the public and private sides of the ledger.

On the public side, some policies entail higher implementation costs than others do, but it is only the non-transfer portion of these costs that matters for social welfare. Thus, while the Authors correctly recognize that nudges often require low (non-transfer) implementation costs (Tor, 2023), this is also the case with traditional financial instruments, like taxes or subsidies, which are mostly resource transfers whose implementation entails some non-transfer costs. Consequently, though implementation costs are sometimes significant for regulatory assessment, their impact on the relative cost-effectiveness and, even more, on the net benefits of non-transfer interventions, like the energy conservation policies evaluated here, is limited. This is not the case, however, for transfer regulation, such as retirement savings instruments, whose CE ratios depend on their non-transfer implementation costs and whose welfare effects cannot be ascertained with standard CBA.

At the same time, the most significant cost categories for other regulations are their private costs generally and their opportunity costs in particular (Tor, 2023). These private costs are largely a function of policy effectiveness, as was the case here with the consumer costs or retailer net revenue losses of the energy conservation policies. Hence, the private costs of more effective interventions tend to comprise a larger fraction of their overall costs, so that their CE performance primarily depends on the private costs they generate for each unit of effectiveness, even while their net benefits depend on the absolute magnitude of these costs compared to policy benefits.

Interestingly, these observations seem to apply even to interventions motivated by public welfare goals, like the energy conservation policies examined here. Besides their (non-transfer) implementation costs, such instruments generate not only private benefits and costs to energy consumers and retailers, but also externality reduction benefits. Except for their implementation costs, however, all of these benefits and costs are a function of policy effectiveness: The more effective an energy conservation policy, the higher its consumer benefits and costs (from retail savings and the welfare losses associated with reduced consumption, respectively), retailer net revenue losses (from reduced sales), and externality reduction benefits (from reduced energy production). Thus, it is the relative magnitudes of these different benefits and costs that matters for overall policy efficiency: When the retailer's net revenue loss is of the same order as the externality reduction benefit, for instance, net policy benefits largely depend on the balance of net consumer welfare versus (non-transfer) implementation costs (cf. Allcott and Kessler, 2019).

Nudges may possess some advantage with respect to non-transfer implementation costs, and occasionally perhaps even with net consumer welfare, when traditional instruments are not particularly effective (e.g., Ito, 2015). However, this potential nudge advantage is of limited generality, since recent evidence indicates that, in the field, many successful nudges exhibit rather small effect sizes (DellaVigna and Linos, 2020).

Our findings show that the evidence examined by Benartzi et al. (2017) does not lead to a general normative prescription for governments to invest more in nudging. Nudges in general are not much more cost-effective than traditional financial instruments, if at all, though some behavioral policies are likely capable of producing some net social benefits. Governments should invest more only in the latter kind of nudges, and only when they generate higher net benefits than competing traditional interventions. Whether and when such particularly net-beneficial nudges are available to policy makers, however, are questions that cannot be answered by the method employed by the Article, only by the same case-by-case CBA that traditional regulations must face.

We therefore agree with the Authors' call for the further study of nudges' welfare effects, which scholars have started to undertake in recent years, but caution that governments or other organizations that aim to promote overall social welfare should not hasten to divert public resources from traditional interventions to behavioral instruments before confirming to the extent possible in each case that the latter can produce the highest net benefits of all available policy alternatives.

Bibliography

Adler, Matthew D., and Eric A. Posner. 2001. *Cost-Benefit Analysis: Legal, Economic, and Philosophical Perspectives.* The University of Chicago Press.

Adler, Matthew D., and Eric A. Posner. 2006. *New Foundations of Cost-Benefit Analysis.* Harvard University Press.

Allcott, Hunt. 2011. "Social Norms and Energy Conservation," 95 *Journal of Public Economics* 1082–1095.

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

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioural Interventions: Experimental Evidence from Energy Conservation," 104 *American Economic Review* 3003-3037.

Allcott, Hunt, and Judd Kessler. 2019. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons," 11 *American Economic Journal: Applied Economics* 236–276.

Arimura, Toshi, Shanjun Li, Richard Newell, and Karen Palmer. 2012. "Cost-Effectiveness of Electricity Energy Efficiency Programs," 33 *Energy Journal* 63–99.

Asensio, Omar I., and Magali A. Delmas. 2015. "Nonprice Incentives and Energy Conservation," *Proceedings of the National Academy of Sciences* E510-E515.

Baskette, C., B. Horiia, E. Kollmanb, and S. Price. 2006. "Avoided Cost Estimation and Post-Reform Funding Allocation for California's Energy Efficiency Programs," 31 *Energy* 1084–1099.

Behavioral Insights Team. 2019. *Annual Report 2017–2018*. https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf.

Benartzi, Shlomo, John Beshears, Katherine Milkman, Cass Sunstein, Richard Thaler, Maya Shankar, Will Tucker-Ray, William Congdon, and Steven Galing. 2017. "Should Governments Invest More in Nudging?," 28 *Psychological Science* 1041–1055.

Boardman, Anthony, David Greenberg, Aidan Vining, and David Weimer. 2018. *Cost-Benefit Analysis: Concepts and Practice*, 5th ed. Cambridge, UK: Cambridge University Press.

Brandon, Alec, John A. List, Robert D. Metcalfe, Michael K. Price, and Florian Rundhammer. 2019. "Testing for Crowd Out in Social Nudges: Evidence from a Natural Field Experiment in the Market for Electricity," 116 *Proceedings of the National Academy of Sciences* 5293-5298.

Carroll, Gabriel, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions," 124 *The Quarterly Journal of Economics* 1639–1674.

Chetty, Raj, John Friedman, Soren Leth-Petersen, and Torben Nielsen. 2014. "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark," *The Quarterly Journal of Economics* 1141-1219.

De Francesco, Fabrizio. 2012. "Diffusion of Regulatory Impact Analysis Among OECD and EU Member States," 45 *Comparative Political Studies* 1277-1305.

De Jonge, Patricia, Marcel Zeelenberg, and Peter W.J. Verlegh. 2018."Putting the Public Back in Behavioural Public Policy," 2 *Behavioural Public Policy* 218-226.

DellaVigna, Stefano, and Elizabeth Linos. 2020. "RCTs to Scale: Comprehensive Evidence from Two Nudge Units" SSRN working paper, https://www.nber.org/papers/w27594.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics* 815-842.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez. 2006. "Savings Incentives for Low- and Middle-Class Families: Evidence from a Field Experiment with H&R Block," 121 *Quarterly Journal of Economics* 1311–1346.

Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez. 2007. "Savings Incentives for Low- and Moderate-Income Families in the United States: Why is the Saver's Credit Not More Effective?," 5 *Journal of the European Economic Association*, 647–661.

Dunlop, Claire, and Claudio Radaelli. 2016. *Handbook of Regulatory Impact Assessment.* London, United Kingdom: Edward Elgar Publishing.

European Commission. 2016. *Behavioral Insights Applied to Policy: European Report 2016.* https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/behavioural-insights-applied-policy-european-report-2016.

European Commission. 2021. *Better Regulation Toolbox.* https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how_en.

Exec. Order No. 12,866, 58 Fed. Reg. 51,735 (Oct. 4, 1993).

Goldin, Jacob, Tatiana Homonoff, Richard Patterson, and William Skimmyhorn. 2020. "How Much to Save? Decision Costs and Retirement Plan Participation," 191 *Journal of Public Economics* 104247.

Hagmann, David, Emily Ho, and George Loewenstein. 2019. "Nudging Out Support for a Carbon Tax," 9 *Nature Climate Change* 484–489.

Hershfield, Hal, Elicia John, Joseph Reiff. 2018. "Using Vividness Interpretations to Improve Financial Decision Making," 5 *Policy Insights from the Behavioral and Brain Sciences* 209-215.

Hitch, Charles Jonston, and Roland N. McKean. 1960. *Economics of Defense in the Nuclear Age*. 1st ed. Harvard University Press.

Ito, Koichiro. 2015. "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program," 7 *American Economic Journal: Economic Policy* 209-237.

Layard, Richard, and Stephen Glaister. 1994. *Cost-Benefit Analysis.* 2nd ed. Cambridge UK: Cambridge University Press.

Levin, Henry, and Clive Belfield. 2015. "Guiding the Development and Use of Cost-Effectiveness Analysis in Education," 8 *Journal of Research on Educational Effectiveness* 400-418.

Levin, Henry, and Patrick McEwan. 2001. *Cost-Effectiveness Analysis.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Miller, Wilhelmine, Lisa A. Robinson, and Robert S. Lawrence. 2006. *Valuing Health for Regulatory Cost-Effectiveness Analysis*. Washington, DC: The National Academies Press.

Minor, Liza. 2019. "DSM Budget Trends Through 2020," *E-Source* (June 14, 2019). https://www.esource.com/429191abpd/dsm-budget-trends-through-2020.

Organisation for Economic Co-operation and Development. 2017. *Behavioral Insights and Public Policy: Lessons from Around the World*. Paris, FR: OECD Publishing. https://read.oecd-ilibrary.org/governance/behavioural-insights-and-public-policy_9789264270480-en.

Organisation for Economic Co-operation and Development. 2018. *Cost-Benefit Analysis and the Environment: Further Developments and Policy Use*. Paris, FR: OECD Publishing. https://www.oecd.org/governance/cost-benefit-analysis-and-the-environment-9789264085169-en.htm.

Organisation for Economic Co-operation and Development. 2020. *Regulatory Policy and COVID-19: Behavioural Insights for Fast-paced Decision Making*. Paris, FR: OECD Publishing. https://read.oecd-ilibrary.org/view/?ref=137_137590-2p5x0tveyp&title=Regulatory-policy-and-COVID-19-Behavioural-insights-for-fast-paced-decision-making.

Pearce, D. W. 1983. *Cost-Benefit Analysis.* 2nd ed. Palgrave Macmillan.

Posner, Eric A. 2003. "Transfer Regulations and Cost-Effectiveness Analysis," 53 *Duke Law Journal* 1067-1110.

Sovacool, Benjamin K., Jinsoo Kim, and Minyoung Yang. 2021. "The Hidden Costs of Energy and Mobility: A Global Meta-Analysis and Research Synthesis of Electricity and Transport Externalities," 72 *Energy Research and Social Science* 101885.

Sugden, Robert, and Alan H. Williams. 1978. *The Principles of Practical Cost-Benefit Analysis*. Oxford University Press.

Tannenbaum, David, Craig R. Fox, and Todd Rogers. 2017. "On the Misplaced Politics of Behavioural Policy Interventions," 1 *Nature Human Behavior* 0130.

Tor, Avishalom. 2020. "The Target Opportunity Costs of Successful Nudges," in Klaus Mathis and Avishalom Tor, eds. *Consumer Law and Economics.* Cham, Switzerland: Springer.

Tor, Avishalom. *Forthcoming 2023*. "The Private Costs of Behavioral Interventions," *Duke Law Journal*.

Tor, Avishalom. 2022. "The Law and Economics of Behavioral Regulation," *Review of Law and Economics*.

Weimer, David. 2017. *Behavioral Economics for Cost-Benefit Analysis: Benefit Validity When Sovereign Consumers Seem to Make Mistakes*. Cambridge, UK: Cambridge University Press.

Wirtshafter Associates. 2006. "Evaluation of the California Statewide 20/20 Demand Reduction Programs."

TABLE 1: ILLUSTRATIVE CORRECTED COST-EFFECTIVENESS OF POLICIES TARGETING ENERGY CONSERVATION[a]

Panel A: Uncorrected Cost-Effectiveness (excluding private costs and including economic transfers)

Article	Policy Type	Instrument	Implementation Costs (per customer)	Effectiveness (average usage reduction per customer)	Uncorrected CE (per kWh saved)
Allcott (2011)	Nudge	Social Information	\$9.72	226.3 kWh	\$0.04
Asensio and Delmas (2015)	Nudge	Social, Environmental and Health Information	\$3,532.23	149.8 kWh[b]	\$23.58
Ito (2015)	Financial Incentive	Conditional Rebate	Overall, \$5.96 Coastal, \$5.81 Inland, \$7.60	Overall, 17.34 kWh Coastal, 3.11 kWh Inland, 169.15 kWh	Overall, \$0.34 Coastal, \$1.87 Inland, \$0.04
Arimura et al. (2012)	Financial Incentive and Education	Demand-Side Management (DSM) Programs	\$12.67	152.54 kWh[c]	\$0.08

[a] All figures updated to 2021 dollars.

[b] Following Benartzi et al.'s (2017) extrapolation from Asensio and Delmas's (2015) data.

[c] Following Benartzi et al.'s (2017) analysis of Arimura et al.'s (2012) results.

Panel B: Corrected Cost-Effectiveness (including private costs and excluding transfers)

Article	Implementation Costs (per customer)	Effectiveness (average usage reduction per customer)	Consumer Costs [*CCs*] (per customer)	Retailer Net Revenue Loss[d] [*RNRL*] (per customer)	Corrected CE I [*including CCs and RNRL*] (per kWh saved)	Non-Transfer Implementation Costs (per customer)	Corrected CE II [*also excluding transfers*] (per kWh saved)
Allcott (2011)	$9.72	226.3 kWh	$2.46[e] or $14.60[f]	$8.49	$0.09 or $0.14	$9.72	$0.09 or $0.14
Asensio and Delmas (2015)	$3,532.23	149.8 kWh	$2.46[e] or $10.31[f]	$13.90	$23.69 or $23.74	$3,532.23	$23.69 or $23.74
Ito (2015)	Overall, $5.96 Coastal, $5.81 Inland, $7.60	Overall, 17.34 kWh Coastal, 3.11 kWh Inland, 169.15 kWh	Overall, $4.17 Coastal, $4.07 Inland, $5.32	Overall, $1.91 Coastal, $0.34 Inland, $18.64	Overall, $0.69 Coastal, $3.29 Inland, $0.19	$1.63	Overall, $0.44 Coastal, $1.94 Inland, $0.15
Arimura et al. (2012)	$12.67	152.54 kWh	$8.87	$3.17	$0.16	$3.42	$0.10

[d] Retailer Net Revenue Loss (RNRL) is the difference between retail electricity cost savings and the retailer's avoidable operating costs (primarily energy acquisition costs).

[e] Using Allcott and Kessler's (2019) estimates of the dollar value of consumer costs of natural gas HERs over one heating season (243 days), and assuming that the consumer costs of annual electricity HERs (or similar instruments) are primarily their direct costs that do not vary with policy effectiveness.

[f] Using Allcott and Kessler's (2019) estimates of the proportion of consumer costs to consumer benefits of natural gas HERs, assuming that the consumer costs of annual electricity use HERs (or similar instruments) include not only direct costs but also the opportunity costs of energy saving behavior that vary with policy effectiveness.

TABLE 2: ILLUSTRATIVE CORRECTED COST-EFFECTIVENESS OF POLICIES TARGETING RETIREMENT SAVINGS[a]

Article	Policy Type	Instrument	Implementation Costs (per affected individual)	Effectiveness (change in annual contributions per affected individual)	Uncorrected CE (per $1 change in contributions)	Resource Transfer	Non-Transfer Implementation Costs (per affected individual)	Corrected CE (excluding transfers) (per $1 change in contributions)
Carroll et al. (2009)	Nudge	Required choice	$2.34	$234.00	$0.01	No	$2.34	$0.01
Chetty et al. (2014)	Financial Incentive	Tax subsidy removal	($228.15)	($631.80)	$0.36	Yes	$2.00	$0.003
Duflo and Saez (2003)	Financial Incentive / Education	Small incentive to attend benefits fair	$4.73	$68.97	$0.07	Yes	$1.00	$0.01
Duflo et al. (2006)	Financial Incentive / Education	20% matching contribution	$19.54	$109.51	$0.18	Yes	$5.00	$0.05
		50% matching contribution	$96.41	$286.07	$0.34	Yes	$5.00	$0.02
Duflo et al. (2007)	Financial Incentive	Tax credit	$10.85	$13.56	$0.80	Yes	$0.50	$0.04

[a] All figures updated to 2021 dollars.

TABLE 3: ILLUSTRATIVE COST-BENEFIT ANALYSIS OF POLICIES TARGETING ENERGY CONSERVATION[a]

Panel A: Net Non-Consumer Welfare (per customer)

Article	Policy Type	Non-Transfer Implementation Costs	Externality Reduction Benefits	Retailer Net Revenue Loss	Net Non-Consumer Welfare
Allcott (2011)	Nudge	$9.72	$7.47	$8.49	($10.74)
Asensio and Delmas (2015)	Nudge	$3,532.23	$4.94	$13.90	($3,541.19)
Ito (2015)	Financial Incentive	$1.63	Overall, $0.57 Coastal, $0.10 Inland, $5.58	Overall, $1.91 Coastal, $0.34 Inland, $18.64	Overall, ($2.97) Coastal, ($1.87) Inland, ($14.69)
Arimura et al. (2012)	Financial Incentive and Education	$3.42	$5.03	$3.17	($1.56)

[a] All figures estimate annual effects per customer, updated to 2021 dollars.

Panel B: Consumer Welfare and Net Social Welfare (per customer)

Article	Consumer Retail Electricity Savings	Consumer Costs	Net Consumer Welfare	Net Non-Consumer Welfare	Net Social Welfare
Allcott (2011)	$33.95	$2.46 or $14.60	$31.49 or $19.35	($10.74)	$20.75 or $8.61
Asensio and Delmas (2015)	$23.97	$2.46 or $10.31	$21.51 or $13.66	($3,541.19)	($3,519.68) or ($3,524.43)
Ito (2015)	Overall, $3.29 Coastal, $0.59 Inland, $32.14	Overall, $4.17 Coastal, $4.07 Inland, $5.32	Overall, ($0.88) Coastal, ($3.48) Inland, $26.82	Overall, ($2.97) Coastal, ($1.87) Inland, ($14.69)	Overall, ($3.85) Coastal, ($5.35) Inland, $12.13
Arimura et al. (2012)	$20.37	$8.87	$11. 50	($1.56)	$9.94