CESifo Working Papers

12775
2026

June 2026

The Value of Behavioral Policies

John A. List, Matthias Rodemeier, Sutanuka Roy, Gregory Sun

CESifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
https://www.cesifo.org

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge
· from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
· from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
· from the RePEc website: https://ideas.repec.org/s/ces/ceswps.html

The Value of Behavioral Policies

John A. List, Matthias Rodemeier, Sutanuka Roy, Gregory Sun

June 16, 2026

Abstract

Behavioral interventions have become central to modern public policy, but their empirical promise remains contested because estimated treatment effects often appear small. We argue that a policy response is economically meaningful only relative to the response generated by alternative policies. We assemble more than 1,200 estimates from over 600 studies comparing "nudges" and traditional price interventions in the markets for cigarettes, alcohol, influenza vaccination, electricity, and residential water. Translating nudge effects into equivalent price changes, we find that behavioral interventions often correspond to enormous fiscal interventions, from an 11% tax on electricity to a 100% subsidy on influenza vaccinations. Nudges are also more cost-effective than price instruments in all markets, but cost-effectiveness does not predict the welfare ranking of policies. Using a behavioral extension of the Marginal Value of Public Funds, we show that nudges have high welfare returns *at the margin*, while price instruments often generate larger *total* surplus at scale.

JEL Codes: D61, D83, H21, I12, I18, Q41, Q48

Keywords: nudges, paternalistic taxes, welfare, energy efficiency, water conservation, influenza vaccination, cigarettes, alcohol consumption, RCTs

Acknowledgments: We thank numerous scholars and seminar participants for helpful discussions. For outstanding research assistance, we thank David Franks, Silvia Amalia Meneghesso, Sara Bernadoni, Massimo Biavardi, and Roberto Fani. Some of the data in this paper were collected in an earlier working paper called "Judging Nudging: Understanding the Welfare Effects of Nudges Versus Taxes". The present paper collects new data, expands the analysis to five markets, and develops a new theoretical framework. Affiliations: List: Dept. of Economics, University of Chicago, NBER, and Australian National University. Rodemeier: Bocconi University, IGIER, BAFFI, CESifo. Roy: Australian National University. Sun: Olin Business School, Washington University.

1 Introduction

Behavioral interventions have become a ubiquitous instrument of public policy. Governments in more than 200 jurisdictions worldwide have established behavioral insights units, deploying reminders, default rules, social comparisons, and informational disclosures across domains ranging from tax compliance and climate policy to public health. The accumulated evidence indicates that these interventions reliably move behavior, although the average treatment effects often seem small in magnitude, especially once taken to scale (DellaVigna and Linos 2022).

We argue that this reduced-form evidence alone cannot answer whether such effect sizes constitute a small or large policy success. Consider a behavioral intervention that reduces energy consumption by 3 percent. Is this a meaningful policy outcome? The answer must depend on what an alternative instrument could achieve in the same market. If demand is very price-inelastic, a 3 percent response could correspond to the demand effect of a 20 percent tax increase, *an enormous fiscal intervention* that would face strong political resistance. Reduced-form effect sizes acquire normative content only when expressed in units of opportunity cost relative to the policy alternative they are meant to replace or complement. This also speaks to the recent criticism that policymakers gravitate toward light-touch behavioral interventions for society's biggest challenges, such as climate change and chronic disease, that may require stricter price-based regulation (Connolly, Loewenstein and Chater 2025). Judging the force of this critique requires translating the nudge into the size of the price instrument it substitutes for.

In this paper, we systematically compare effect sizes of nudges and taxes within the same markets based on hundreds of estimated effects. We assemble a hand-coded database of more than 1,200 causal point estimates from over 600 studies, spanning two policy instruments in five consumer markets: cigarettes, alcohol, influenza vaccines, household electricity, and residential water. To our knowledge, no prior dataset combines the systematic coverage of behavioral interventions in the spirit of DellaVigna and Linos (2022) with parallel meta-analytic coverage of price interventions in the same markets. We estimate a simple statistic that allows policymakers to judge the economic magnitude of nudges: The *equivalent price change* (EPC), which is the ratio of the average nudge treatment effect to the average price elasticity in a market. This statistic tells us how large the corresponding fiscal instrument would need to be in order to generate the same effect as the nudge—an intuitive indicator of the intervention's policy scale and opportunity cost.

Two key findings emerge. First, behavioral interventions generate demand responses comparable to *substantial price interventions*. The average electricity conservation nudge has an EPC of 11 percent, meaning policymakers would need to impose an 11% tax on electricity to replicate the nudge effect—a tax policy that would be highly unpopular politically and difficult to implement. The cigarette, water and alcohol nudges have EPCs of 28, 17 and 34 percent, respectively. All of them imply large tax changes would be necessary to mimic their effects. Most strikingly, the average flu shot nudge has an EPC of around 100 percent: current vaccine nudges replicate the demand response of a subsidy that fully covers the vaccine price. Behavioral interventions are doing the work of non-trivial fiscal interventions.

Second, nudges are uniformly more cost-effective than price instruments at changing behavior, but *cost-effectiveness does not provide the correct welfare ranking of policies*. Since nudges are often cheap to implement, they typically look very attractive in terms of cost-effectiveness. This echoes early evidence in Benartzi et al. (2017) for a smaller set of interventions. In our data, nudges generate more behavioral change per dollar than price instruments in all five markets.[1] Yet, cost effectiveness may not be

[1]The EPC overcomes limitations of cost-effectiveness as a comparative policy statistic. First, only

a good proxy for welfare. We follow Hendren and Sprung-Keyser (2020) in adopting the Marginal Value of Public Funds as the welfare object and extend their framework to settings where consumers may fail to maximize utility. The extension requires an assumption about how nudges map into utility, a question most studies do not pin down. Here, we allow for model ambiguity and test different assumptions. Our main specification follows the reduced-form approach to behavioral public finance (Mullainathan et al. 2012) in assuming that nudges reduce choice mistakes (e.g. by lowering forgetfulness to get a flu shot, reducing inattention to energy costs, or improving knowledge about the health consequences of smoking). In this case, $EPC \times price$ is part of the money-metric welfare gains to consumers from being nudged. We relax this assumption in robustness exercises but adopt it as our leading case due to its ubiquity in the literature where the EPC and closely-related ratios have been used to identify welfare (e.g., Chetty et al. 2009, Bertrand et al. 2010, Ferraro and Price 2013, Taubinsky and Rees-Jones 2017, Dubois, Griffith and O'Connell 2018, Houde 2018*a*, Allcott and Kessler 2019, Beshears et al. 2020, Houde and Myers 2021, Houde 2022, Rodemeier and Löschel 2025, and many others). Outside behavioral economics, EPC-style ratios have an even longer tradition of identifying marginal utility (e.g., Berry, Levinsohn and Pakes 1995).

The resulting welfare estimates yield a more nuanced picture than cost-effectiveness alone. In line with cost-effectiveness, nudges often generate large welfare gains *per public dollar*. Energy, water, alcohol, and cigarette nudges have MVPFs of 2.0, 2.3, 4.4, and 20, and vaccine nudges an MVPF of 62. These values compare favorably with many conventional social programs, documented in Hendren and Sprung-Keyser (2020), including cash transfer programs, housing vouchers, and unemployment insurance. But

subsidies, but not taxes, have a meaningful cost-effectiveness, because taxes *generate* rather than expend revenue. The EPC sidesteps this asymmetry. Second, the EPC conveys political feasibility directly. A 5 percent EPC for cigarettes corresponds to a tax change well within the range jurisdictions routinely enact; a 40 percent EPC for household electricity corresponds to one that would face mass demonstrations by the public. Third, the EPC is a key input for any welfare analysis as we explain further below.

per-dollar performance differs from the *total welfare gains at scale*. We compute each policy's welfare as the dollar-valued surplus. This metric can rank policies differently because the total gain is bounded by the size of the behavioral response each instrument can produce. Nudges produce only the modest effects documented in our meta-analysis, whereas taxes and subsidies can generate substantially larger ones. Some markets exhibit failures large enough that only price instruments can close the welfare gap. In energy, the optimal corrective tax generates roughly *seven times* the total welfare of the optimal nudge, even though the nudge is more cost-effective. Similar patterns arise in influenza vaccines, water, and alcohol. Cigarettes are the only market in our sample where nudges continue to dominate taxes even when scaled up. The broader pattern is unambiguous: at scale, price instruments dominate nudges in most markets, and cost-effectiveness does not provide the correct welfare ranking of policies.

The policy implication is that nudges often deliver large effects at low fiscal cost, and they can be powerful complements to taxes and subsidies. But they rarely eliminate the need for price-based policy at scale.[2]

We test the robustness of our results to alternative normative assumptions, such as that nudges may not fully improve choice quality. While this changes the absolute value of the welfare effects of nudges, our *qualitative* policy rankings remain unchanged. This implies that—despite the normative ambiguity about how nudges map into welfare—meta-analytic approaches in the spirit of Hendren and Sprung-Keyser (2020) can still provide useful policy guidance.

Our paper contributes to three distinct literatures. First, we extend the MVPF framework to accommodate behavioral biases, an extension previously called for by Finkelstein

[2]This finding speaks directly to concerns that governments may use nudges as politically convenient substitutes for stronger regulation (Connolly, Loewenstein and Chater 2025). Our estimates suggest that this concern is warranted in several markets: choosing nudges over price instruments can leave substantial welfare gains unrealized.

and Hendren (2020). The extension allows for the possibility that consumers make mistakes and public policies to act directly on those mistakes (see, e.g., Farhi and Gabaix 2020). While previous research has conducted behavioral welfare analysis within specific contexts using experimental or structural methods, these studies are often limited to a specific intervention type or market (e.g., Dubois et al. 2018, Houde 2018*a*, Houde 2018*b*, Allcott, Cohen, Morrison and Taubinsky 2025, Rodemeier 2025, and others). We complement this literature by aggregating these insights into a broader bird's-eye view and evaluating welfare effects across multiple interventions and markets. Second, from an applied policy perspective, our results provide a comprehensive overview of the effects of nudges and taxes in health and environmental policy (e.g., Milkman et al. 2011, Allcott and Kessler 2019, Myers and Souza 2020, Andor et al. 2020, Campos-Mercade et al. 2021, Schneider et al. 2023, and others).[3][4]

Third, our paper offers a novel database of reduced-form effects of behavioral interventions and relates to a growing literature using meta analyses on nudging (Benartzi et al. 2017, Antinyan and Asatryan 2019, Hummel and Maedche 2019, DellaVigna and Linos 2022). Different from prior work, however, we leverage reduced-form estimates of both non-price and price interventions to draw insights on policy rankings. We further highlight how methodologies in this spirit can provide deep implications for the choice of the optimal policy tool.

The remainder of our study is structured as follows. Section 2 presents our meta-analysis of nudge and price elasticities. Section 3 shows how these results can be used to estimate a behavioral version of the MVPF. Here we compare welfare effects of each policy tool within and across markets. We conclude in Section 4.

[3]See Rodemeier and Semken (2026) for a recent review of behavioral interventions in energy & environmental economics.

[4]This also relates to Hahn et al. (2024), who use the MVPF framework to compare climate-related tax and spending policies under the assumption that consumers are fully rational.

2 Meta-Analysis of Reduced-Form Effects

We document the reduced-form benchmark that the paper is built on: average nudge effects, average price elasticities, and the equivalent price change that connects them. These are the primary empirical objects because they let us compare non-price and price interventions in a common behavioral unit, without committing yet to a particular welfare model. The welfare analysis later will translate that benchmark into normative language under explicit assumptions.

2.1 Inclusion Criteria

Our empirical focus relies on using results from consumer markets where behavioral interventions and price elasticities are well-studied. We found that data availability is sufficient to support a quantitative meta-analysis in five markets: cigarettes, alcohol, influenza vaccines, household electricity, and residential water. To consistently compare effects across studies, we must choose standardized outcome variables. For cigarettes, we included all papers addressing cigarette usage. The most commonly used outcome of interest in studies is the probability to *stop* smoking, i.e., a form of extensive margin. However, we also have a number of studies that document both the cessation probability *and* changes in number of cigarettes smoked (i.e., the intensive margin) for each subject. We use this information to estimate the mapping between cessation elasticity and aggregate cigarette demand. For influenza vaccinations, our outcome of interest is vaccination probability. We include only papers that address seasonal flu vaccination, not other epidemic strains. For household electricity consumption, we include all papers measuring residential electricity use. For water, our outcome is household water consumption, typically measured through utility billing records. For alcohol, we include studies reporting changes in measured ethanol consumption.

For comparison purposes, we also need to convert point estimates and standard errors of absolute treatment effects into percent changes. This conversion required papers to report specific information. For the cigarette and vaccine markets, this was straightforward since the outcome variables were binary, so means and sample sizes within different treatment groups suffice to characterize the asymptotic distribution of estimates. In general, all included studies had to report either a standard error or enough information to impute one. We excluded any paper where this was not possible.

Definition of Nudges In addition to the selection of markets, we must define what constitute "price interventions" and "nudges". Our definition of price intervention is relatively straightforward, including any policy that changes the sales price of the underlying good. By contrast, precisely defining a nudge is harder. For instance, one might argue that any non-price intervention is a nudge. However, such a definition would also imply that medication or therapy with a psychologist (e.g., to reduce nicotine addiction) would be included in our analysis. Such an example would be inconsistent with how the literature typically understands nudges. To remain consistent with the literature, we follow the spirit of Benartzi et al. (2017) in choosing which interventions are nudges. We exclude financial incentives, legal mandates, medication, and therapy sessions but include "informational" nudges (e.g., giving consumers information about consumption of peers, effects on health, or billing), "planning" nudges (e.g., goal-setting prompts and offering consumers the ability to pre-commit to choices), and "streamlining" nudges that simplify tasks (e.g., setting defaults and pre-selecting a choice set). We deviate from Benartzi et al. (2017) by counting "educational interventions" as nudges. We include these interventions because there is no clear distinction between what constitutes information provision (a form of nudge in Benartzi et al. 2017) and an educational intervention. For instance, a leaflet describing health risks from smoking may be defined as either. To avoid this ambiguity, we therefore classified any informational intervention as a nudge.

2.2 Data

Data on nudges were obtained primarily by scraping Microsoft Academic search results. We chose Microsoft Academic because it was easy to scrape and had similarly good coverage as other potential sources (Martín-Martín et al. 2021). Our scraping procedure yielded titles, reference information, abstracts, and lists of citations and references. Our data collection took place prior to the large-scale adoption of Artificial Intelligence and all data was hand-coded manually by research assistants.[5]

Most papers related to nudging influenza and smoking are from the medical literature. The digitization rate of the medical literature, including older literature is fairly high. For the literature on energy nudges, we found that this scraping approach missed relevant papers. We therefore consulted an additional meta-analysis by Delmas, Fischlein and Asensio (2013) as well as another survey of the literature by Darby et al. (2006). We searched through the studies that were cited or mentioned in these two additional sources, and added any that were missed from scraping. For water nudges, we followed a parallel procedure, supplementing the scraped records with studies referenced in reviews and empirical work on behavioral water conservation. The sample includes studies mainly of field experiments on residential water conservation. For alcohol nudges, the majority of studies originate in the medical and public health literatures on brief interventions, which we retrieved via the same systematic search procedure used for cigarettes and influenza vaccines.

[5]We assigned two research assistants (RAs), working independently, to check the first 500 results for each search term from the scraping outcomes; first by title, then by abstract. We reconciled the works of the RAs by checking them against each other. Two RAs then read each paper and, if it was deemed fit for inclusion, independently collected the relevant information—treatment effects, sample sizes, p-values and other statistics, as well as information on the details of the intervention.

There are a variety of high-quality meta-analyses estimating price elasticities. For cigarette price elasticities, we obtained the underlying data of the meta-analysis by Gallet and List (2003). For energy price elasticities, we collected all the papers mentioned in Zhu et al. (2018), excluding all estimates of price elasticity not meeting our inclusion criteria. We further obtained data from Alt et al. (2024) which provided studies that included price and nudge interventions within the same sample. For water and alcohol price elasticities, we draw on a large empirical literature and obtained underlying data in Jegnie et al. (2021) and Nelson (2013), respectively. We were unable to find a meta-analysis for price elasticities of influenza vaccines. We therefore searched Microsoft Academic and Google Scholar for all papers estimating the effects of monetary incentives on vaccination take-up. We also estimated price elasticities from policies that made flu shots free to subjects. Because of the difficulties in finding studies on price interventions in the vaccine market, the sample of studies we were able to obtain in this market is smaller than in the other four markets.

Overall, we gathered 1,289 point estimates from 610 studies on nudges and price interventions across these five markets.

2.3 Reduced-Form Effects

Figure 1 presents the distribution of nudge effects (Panel A) and of price elasticities (Panel B) in each market. Energy nudges reduce household electricity consumption by, on average, 4.75% ($p < 0.001$), with a median of 4.00% and a standard deviation of 5.18%. Large reductions above 10% are unlikely. These interventions are dominated by Home Energy Reports that compare a household's consumption to that of similar neighbors, as well as real-time usage feedback and tips on conservation.

The average treatment effect of smoking nudges constitutes a 7.5% increase in the cessation probability ($p < 0.001$). The median effect is 5.96%. Using studies that map the cessation probability into aggregate demand for cigarettes, we find that this 7.5% increase in the cessation probability amounts to an approximate 14% reduction in cigarette demand. We also observe a large amount of heterogeneity in effects: The standard deviation is 11.71% and large values, such as a reduction in the smoking probability of 30%, are possible. The interventions vary widely, ranging from warning labels and graphic images on cigarette packs to planning prompts that pre-commit smokers to a quit date.

For influenza vaccines, we observe the largest relative effects across markets. Here, the average probability to get a flu shot increases by 34.79% due to nudges, the median is 18.38%. The standard deviation is 39.70%. Interventions in this market are mostly reminders sent by text, email, or letter, opt-out scheduling of vaccination appointments, and information about the risks of influenza and the benefits of immunization.

In the market for residential water, nudges reduce consumption by 6.53% on average, with a standard deviation of 7.26%. Effects are driven primarily by social norm comparisons, feedback on usage, and informational interventions.

For alcohol consumption, the average nudge reduces consumption by 19.10%, with a standard deviation of 11.87%. The interventions in the alcohol literature are predominantly informational and goal setting interventions.

Turning to Panel B, the own-price elasticity of demand is 0.45, 0.49, 0.33, 0.38, and 0.57 in the markets for energy, cigarettes, flu vaccines, water, and alcohol, with substantial heterogeneity within each market. The underlying literatures vary widely in size: 68 estimates for electricity, 94 for cigarettes, 174 for residential water, and 560 for alcohol, but only 9 for influenza vaccines. The influenza vaccine literature is small because direct price variation on vaccines is rare, partly reflecting concerns that paying people to get vaccinated is ethically contested and could backfire. The estimates we do have show no such backfire: price reductions always raise vaccination take-up, consistent with recent evidence on COVID-19 vaccine incentives (Schneider et al. 2023).

The averages reported above mask substantial heterogeneity across nudge subtypes. Appendix A reports the most effective nudge categories in each market: social comparisons in energy, information and feedback interventions in cigarettes, reminders and defaults in flu vaccines, commitment and goal-setting in water, and social-norm and feedback interventions in alcohol.

Funnel-plot diagnostics in Appendix B indicate publication bias for the energy, cigarette, and flu vaccine treatment effect estimates (Egger test $p < 0.05$), and no significant evidence of bias for water or alcohol. In the same appendix, we show that our key results, like the EPC and welfare rankings of policies, are robust to correcting for publication bias.

2.4 The Equivalent Price Change of Nudges

To compare nudges and price instruments on a common footing, we ask how large a tax or subsidy change would be needed to reproduce the same first-order behavioral response as the nudge. We formalize this as the equivalent price change (EPC) of a nudge, defined as the ratio of the average nudge treatment effect (expressed as a percent of baseline levels) to the average price elasticity in a given market. Figure 2 presents the EPC in each market. Energy nudges have an average effect that is equal to a 10.7% tax on the household electricity price. This is a substantial effect, which would amount to 1.41 cents per kWh tax on the average US electricity price.

The average smoking nudge induces the same demand response as a 28.1% sales tax on the cigarette price, statistically significant at the 1% level.[6] In the US, this implies a tax of roughly $2.25 per cigarette pack, which is close to the current average tax rate in

[6] The price elasticity refers to continuous cigarette demand rather than the cessation probability, so we express the smoking nudge effect in the same units.

the US of \$3.01/pack.

Influenza vaccines have an extremely high price change equivalent: on average they cause the same increase in vaccination rates that a 104.2% subsidy on the vaccine price would induce. Despite the few price studies we saw in Figure 1, we are sufficiently powered to conclude that this effect is statistically different from zero ($p < 0.01$). The magnitude of the effect implies that we would want to make flu shots free to everyone and—if we interpret the coefficient at face value—even offer a small financial payment of 4.2% of the vaccine price to people.

In the water market, the average nudge induces the same response as a 17.3% price increase, reflecting the combination of moderate behavioral responses and a relatively inelastic demand. At a typical US water-and-sewer rate of \$11.50 per hundred cubic feet (CCF), this corresponds to an additional \$2 per CCF, or roughly \$200 per US household per year.

For alcohol, the average nudge is equivalent to a 34.2% price increase. Applied to a typical US retail price, this corresponds to roughly \$0.43 per 10 g of pure ethanol, or about \$0.60 per US standard drink, several times the average combined US federal and state alcohol excise.

In conclusion, behavioral non-price interventions induce substantial demand responses if we compare them directly to the change in taxes and subsidies required to match this effect. This equivalent price change is an important statistic because it moves from reduced-form effects (which are hard to judge normatively) to a statistic that directly relates it to policy alternatives.

2.5 Correlation Between Price and Nudge Effects

An important question for policymakers is whether nudges are complements or substitutes to price instruments. Complementarity means that nudges manage to change the behavior of a subpopulation of consumers that are less likely to respond to price changes. Answering this question is also important for the welfare impact of a nudge: As we discuss in more detail in Section 3, this welfare gain will be particularly positive if nudges help high-value consumers to make better choices. If price elasticities reflect marginal utility of consumption, nudges have particularly large welfare effects if price-inelastic consumers are nudge elastic.

Despite the importance of this correlation, most studies do not feature both a price and nudge intervention. Through targeted searches by our research assistants, we identify 21 usable paired observations, allowing us to estimate the correlation for two markets: cigarettes and energy consumption. While this covers only a subset of the markets of interest, it provides direct evidence for one health market and one resource-consumption market.

In both markets the point estimates are positive, but only the energy correlation is statistically significant: $r = 0.636$ ($p = 0.019$) for energy and $r = 0.315$ ($p = 0.447$) for cigarettes. Consumers whose demand responds more strongly to price changes also exhibit larger behavioral responses to nudges in the energy sample; the cigarette estimate is imprecise. This pattern suggests that households who have relatively weak preferences for the underlying good may also be more likely to change their behavior when nudged, consistent with both neoclassical models and models of bounded rationality or rational inattention.

From a policy perspective, the results imply that nudges are substitutes to taxes as they tend to induce behavioral change among the same consumer types. This becomes important again when we discuss welfare implications in Section 3.

2.6 Cost-Effectiveness

Finally, we follow Benartzi et al. (2017) and compute the cost-effectiveness of each intervention. While nudge implementation costs are typically not reported in each study, we managed to obtain 44 usable cost observations across the five markets. To compute cost-effectiveness, we divide the average treatment effect by the average nudge provision costs.[7]

Before getting to the results, we should note that cost-effectiveness is a natural metric only for *spender* policies. Nudges and subsidies expend public funds, so dividing behavioral change by dollars expended has a transparent interpretation: it tells the planner how much behavior change each dollar of program spending purchases. Corrective taxes do not expend public funds, they collect them, and a dollar of tax revenue is a transfer rather than a resource cost.[8] For the purpose of putting nudges and taxes on a common cost effectiveness scale, we compute the average behavioral response per $1 of *fiscal flow*. Here, flow is program spending for nudges and subsidies, but the tax burden of taxes on consumers. This convention treats a dollar of tax revenue as the cost borne by consumers. Thus, we do not make a distinction here over whether the cost of the policy is financed by the government budget (as is the case for nudges and subsidies) or by the consumer budget. These nuances are better captured in the MVPF framework which we turn to in Section 3.

Figure 4 illustrates the cost-effectiveness of nudges versus price interventions across

[7]For this exercise we do not need to assume that implementation costs and treatment effect size are uncorrelated. But we have to make the assumption that whether studies *report* implementation costs or not is not directly linked to the magnitude of the treatment effect. We acknowledge that this assumption might be violated but view our approach as the best solution given the information provided by the literature.

[8]This is another advantage of our EPC measure that can be computed for both spenders and revenue raisers.

markets. To ensure comparability across markets with different population sizes and consumption patterns, we normalize by distributing a fixed budget of $10 million across the relevant population in each market. The metric displayed is the aggregate percentage behavior change per dollar invested, which accounts for both the per-person effect size and the number of people reached.

The results reveal substantial heterogeneity across markets. In the energy market, nudges reduce consumption by 0.10% per dollar while taxes only by 0.03% per dollar. For cigarettes, nudges lower demand by 1.03% per dollar, substantially outperforming taxes with 0.02% per dollar. Influenza vaccine nudges exhibit the highest cost-effectiveness among all interventions, generating 7.16% aggregate effect per dollar, followed by vaccine subsidies at 1.68% per dollar. Water conservation nudges reduce consumption by 0.46% per dollar versus taxes by 0.09%. For alcohol, nudges attenuate consumption by 0.38% per dollar, compared to 0.11% per dollar for alcohol taxes. The broader pattern across markets suggests that nudges are always more cost-effective at moving aggregate consumption behavior than taxes per dollar invested, echoing the main message in Benartzi et al. (2017).

The open question we aim to address next is whether these cost effectiveness measures are informative about different measures of welfare effects of the interventions.

3 From Cost-Effectiveness to Welfare

Section 2 documented that nudges generate behavioral responses comparable to substantial price interventions and do so at low fiscal cost. Translating these reduced-form facts into welfare requires a framework that values the behavioral response in normative units. To do so, we develop a behavioral extension of the Marginal Value of Public Funds (MVPF, Hendren and Sprung-Keyser 2020). To our knowledge, this is the first MVPF construction that accommodates behavioral biases, although Finkelstein and Hendren (2020) have previously discussed such an extension informally.

3.1 Environment

We consider a market with a unit mass of heterogeneous consumers indexed by i. Each consumer has quasi-linear preferences $u_i(x_i) + y_i$ over a consumption good x_i and a numeraire y_i, which they maximize subject to budget constraint $px_i + y_i \leq M_i$, where $p = c + t$ is the tax-inclusive price and c is the constant marginal cost. We assume $px_i < M_i$ for all i, so demand depends only on price. This is reasonable for our markets where prices are small relative to income. Let $D'_i = \partial x_i / \partial p < 0$ denote consumer i's demand slope.

Two frictions motivate policy intervention. First, each unit of consumption generates a per-unit externality ξ, which we treat as random in the policymaker's information set with prior mean $\mathbb{E}[\xi]$ and variance $\mathrm{Var}(\xi)$, reflecting epistemic uncertainty about external damages. The externality enters the social welfare function but not private choices. Second, consumer i has a behavioral bias b_i (in money-metric terms) that distorts her private first-order condition. Rather than equating marginal utility with price, $u'_i(x^*_i) = p$, she solves

$$u'_i(x_i) = p - b_i \tag{1}$$

so that $b_i > 0$ implies overconsumption (perceived marginal benefit exceeds price) and $b_i < 0$ implies underconsumption. The socially optimal (first-best) allocation satisfies $u'_i(x_i^{FB}) = c + \mathbb{E}[\xi]$ for each i. Preferences and biases may covary arbitrarily. For the marginal externality, we assume $\xi \perp (b_i, D'_i)$ because we think of uncertainty in ξ reflecting the policymaker's imprecise knowledge of external damages (e.g., about the true value of the Social Cost of Carbon) rather than individual-level variation in marginal damages.[9]

The government has two instruments. A *corrective tax* t that raises the consumer price to $c+t$. And a *nudge* $n \in [0,1]$ that is characterized by two additional primitives, a de-biasing intensity θ and a per-unit psychic cost $\kappa \geq 0$. We let $n=0$ denote no nudge and $n=1$ the deployed nudge we observe in the data. Consumer i's perception wedge is $(1-\theta n)b_i$, so the deployed nudge shifts the wedge from b_i to $(1-\theta)b_i$. If $\theta = 1$, the nudge restores rational choices; $0 < \theta < 1$ is a partially corrective nudge. The psychic cost $\kappa n\, x_i$ is proportional to consumption and could, for example, represent the direct disutility from seeing a disturbing image on cigarette packages that highlights cancer risk.

Her modified first-order condition is

$$u_i'(x_i) = p + \kappa n - (1-\theta n)b_i.$$

The reduced-form approach to behavioral public finance, dating to Chetty et al. (2009) and Mullainathan et al. (2012), treats observed treatment effects as evidence of bias correction and identifies welfare from the elasticity-weighted demand response. In our notation this corresponds to $\theta = 1$ and $\kappa = 0$: the nudge closes the bias it was designed to address, with no residual psychic cost. The link to the EPC is direct at the individual level. Because b_i enters consumer i's first-order condition (1) like a shift in the perceived price, a full-debiasing nudge induces a quantity response equal to her demand slope times b_i. Expressing the demand response to a de-biasing nudge as a percent change T_i and dividing by the consumer's own-price elasticity ϵ_i recovers her dollar-denominated bias: $b_i = (T_i/\epsilon_i)\cdot p$. This is the individual-level EPC times the consumer price. With het-

[9]The independence assumption does *not* mean that consumers with different preferences and biases produce the same amount of externalities; it simply means that the externality *per unit of consumption* is the same across consumers.

erogeneous consumers the welfare-relevant population analogue is an average over those marginal to the nudge; we formalize this later in equation (3) below. The EPC introduced in Section 2 is therefore the central input to the welfare analysis.

Similar specifications have been the workhorse of a large empirical literature that uses EPC-style ratios to identify welfare of behavioral interventions (Chetty et al. 2009, Allcott and Kessler 2019, Bertrand et al. 2010, Beshears et al. 2020, Dubois, Griffith and O'Connell 2018, Ferraro and Price 2013, Houde 2018*a*, Houde and Myers 2021, Houde 2022, Rodemeier and Löschel 2025, Taubinsky and Rees-Jones 2017). The revealed-preference logic of scaling a quantity response by the relevant price elasticity to recover marginal utility predates this work in industrial-organization demand estimation (Berry, Levinsohn and Pakes 1995).

We test robustness to the most direct experimental evidence against this approach. Allcott, Cohen, Morrison and Taubinsky (2025) design two experiments to measure how much nudges de-bias, θ, and how much they generate direct disutility, κ. Their experiment designs a normative benchmark for choice quality and observes whether choices move to this benchmark as subjects get nudged. Results suggest that nudges on average close only a fraction of the underlying bias ($\theta < 1$), with essentially zero psychic cost on average ($\kappa = 0$). As a robustness exercise, Appendix D substitutes their estimates into our framework and also test robustness to a large variety of alternative values. While this changes quantitative magnitudes, the qualitative policy rankings are very robust. Only in one of five markets could we construct a scenario in which our policy ranking in the main text changes. Our main conclusions are, therefore, not artifacts of the $\theta = 1$ assumption we proceed with now.

3.2 The Marginal Value of Behavioral Policies

We first adopt the MVPF in Hendren and Sprung-Keyser (2020) as our welfare object, which is the welfare gain per additional dollar of government expenditure:

$$\text{MVPF}_j = \frac{\sum_i WTP_i^j}{G_j}$$

where WTP_i^j is individual i's (normative) willingness to pay for policy j and G_j is the net cost to the government (implementation costs plus fiscal externalities from behavioral responses). For a rational consumer, the envelope theorem implies that a tax change dt reduces utility by $x_i\, dt$ and raises revenue $x_i\, dt + t\frac{\partial x_i}{\partial t}\, dt$. Assuming no existing externalities or taxes, this yields MVPF $= 1$. Different from this established framework, we show that the *behavioral version* of the MVPF departs from one, even absent externalities and pre-existing taxes, because consumers make mistakes and behavioral responses have first-order welfare consequences.

The MVPF has two specific advantages over more traditional welfare analyses that measure changes in excess burden using a utilitarian welfare function. First, the MVPF does not require us to take a stance on distributional preferences of the policymaker. To illustrate, consider a corrective tax "A" with $\text{MVPF}_A = 1$, meaning each dollar of revenue raised costs the taxed group exactly one dollar in willingness to pay, and a nudge "B" with $\text{MVPF}_B = 0.5$, meaning each dollar of public spending generates only fifty cents of willingness to pay for its recipients. Then, the budget-neutral reform that raises one dollar through A and spends it on B improves social welfare if and only if $\eta_B \cdot \text{MVPF}_B > \eta_A \cdot \text{MVPF}_A$ where η_A and η_B are the social marginal welfare weights on the taxed and recipient groups. With our numbers this collapses to $\eta_B/\eta_A > 2$: financing nudge B through corrective tax A is welfare-enhancing if the social planner values the welfare of recipients of B more than twice as much as the welfare of those who pay for the tax.

A second advantage is that the MVPF places nudges and taxes on the same conceptual scale as other public policies evaluated in the MVPF literature, allowing us to compare the social value of behavioral interventions with other policies that governments use in different markets, like social insurance, education programs, health policies, etc.

MVPF of Taxes Consider first a small change in the tax rate. Consumer i's total utility change from a tax increase dt is

$$dU_i = (u_i'(x_i) - p)\,\frac{\partial x_i}{\partial t}\,dt - x_i\,dt = -b_i\,\frac{\partial x_i}{\partial t}\,dt - x_i\,dt$$

where we used $u_i' - p = -b_i$ from equation (1). When the consumer *overvalues* consumption by $b_i > 0$, the welfare-relevant correction from the behavioral response is $-b_i \frac{\partial x_i}{\partial t} > 0$ (positive, because $\frac{\partial x_i}{\partial t} < 0$ and $b_i > 0$): the tax forces the overconsumer to cut back, which improves her normative welfare. To obtain normative willingness to pay, we must consider what a consumer would be willing to pay to the policymaker to correct them, given the consumer was made meta-cognitively aware of her mistakes. In other words, we must quantify the consumer surplus gains from corrective policy.[10]

Adding the externality correction, $-\mathbb{E}[\xi]\frac{\partial x_i}{\partial t}\,dt > 0$, to the bias correction, the consumer's overall normative WTP for the tax is

$$WTP_i = -b_i\,\frac{\partial x_i}{\partial t}\,dt - x_i\,dt - \mathbb{E}[\xi]\,\frac{\partial x_i}{\partial t}\,dt.$$

[10]This logic is analogous to externalities: while consumers ignore externalities in their private consumption choices, their normative willingness to pay to internalize them—as for example implied by political preferences—is positive.

Aggregating across consumers, we have

$$\sum_i WTP_i = \left[-(\mathbb{E}[b]+\mathbb{E}[\xi])\,\mathbb{E}\left[\tfrac{\partial x_i}{\partial t}\right] - \mathrm{Cov}\big(b, \tfrac{\partial x_i}{\partial t}\big) - \mathbb{E}[x_i]\right] dt.$$

At $t = 0$ the net government cost of imposing the tax is negative, $G = -\mathbb{E}[x_i]\,dt$ because the tax *raises revenue* rather than imposes costs. Dividing by G, we obtain

$$\mathrm{MVPF} = 1 + \frac{(\mathbb{E}[b]+\mathbb{E}[\xi])\,\mathbb{E}\left[\tfrac{\partial x_i}{\partial t}\right] + \mathrm{Cov}\big(b, \tfrac{\partial x_i}{\partial t}\big)}{\mathbb{E}[x_i]}.$$

With $\mathbb{E}[b] > 0$, $\mathbb{E}[\xi] > 0$, and $\mathbb{E}[\partial x_i/\partial t] < 0$, the second term is negative, so $\mathrm{MVPF} < 1$. For revenue raisers like taxes, a MVPF below one is desirable: it means that to raise 1 dollar in tax revenue from consumers we forgo less than 1 dollar in consumer welfare. In the absence of behavioral biases ($b = 0$) and externalities ($\mathbb{E}[\xi] = 0$), the MVPF equals one: each dollar of tax revenue imposes exactly one dollar of welfare cost on consumers.

The MVPF of a tax is lower (i.e., the welfare gain from raising revenue is larger) when $\mathrm{Cov}(b, \partial x_i/\partial t) < 0$, so when the bias is concentrated on price-elastic consumers. To understand the economics behind this result, note that for a given magnitude of b, a larger price elasticity causes a larger deviation in quantity consumed from the private optimum.[11] Since marginal utility decreases in consumption, larger deviations from the optimum imply larger forgone consumer surplus. A corrective tax recovers forgone surplus, so its welfare gain is largest among consumers who are both heavily biased and highly price-elastic.

MVPF of Nudges The nudge changes consumer i's allocation by $(\partial x_i/\partial n)\,dn$. The consumer's normative optimality gap is $b_i + \mathbb{E}[\xi]$ (she overconsumes by b_i due to bias and generates externality $\mathbb{E}[\xi]$ per unit in expectation). Each unit of reduced consump-

[11] This becomes obvious from the consumers' FOC where b enters the same way as the price of the good.

tion generates a normative benefit of $b_i + \mathbb{E}[\xi]$, so the WTP per consumer is $-(b_i + \mathbb{E}[\xi])(\partial x_i/\partial n)\, dn > 0$, which is positive when $\partial x_i/\partial n < 0$ for overconsumption goods. The net government cost is $G = C'(n)\, dn - t\, \mathbb{E}[\partial x_i/\partial n]\, dn$, where $C'(n)$ is the marginal implementation cost.

Aggregating and dividing by $G = C'(n)\, dn$, we obtain

$$\text{MVPF}_{\text{nudge}} = \frac{-(\mathbb{E}[b] + \mathbb{E}[\xi])\, \mathbb{E}\big[\frac{\partial x_i}{\partial n}\big] - \text{Cov}\big(b, \frac{\partial x_i}{\partial n}\big)}{C'(n)}.$$

Nudges have larger welfare gains when (i) the expected distortions cause large overconsumption and nudges decrease consumption (i.e., $(\mathbb{E}[b] + \mathbb{E}[\xi]) > 0$ and $\frac{\partial x_i}{\partial n} < 0$), (ii) when they reduce overconsumption more for high-biased consumers ($\text{Cov}\big(b, \frac{\partial x_i}{\partial n}\big) < 0$), and (iii) when implementation costs are low.[12]

3.2.1 Empirical Identification

The MVPF formulas derived above express welfare in terms of unobserved primitives: the bias b_i, the demand slope D_i', and the externality ξ. We now show how each of these objects is identified in our meta-analysis.

The key identification equation follows from the model's first-order condition, aggregated to the population level. Under the de-biasing interpretation ($\kappa = 0$), a nudge of effectiveness θ shifts consumer i's perceived price by θb_i, generating an individual consumption response $\Delta x_i = D_i' \cdot \theta b_i$. Aggregating across consumers and dividing by mean consumption $\bar{x}$ gives the average percent treatment effect

$$\bar{T} \;\equiv\; \frac{\mathbb{E}[\Delta x_i]}{\bar{x}} \;=\; \frac{\theta \cdot \bar{\epsilon} \cdot \mathbb{E}_W[b]}{p}, \tag{2}$$

[12]Conversely, if $(\mathbb{E}[b] + \mathbb{E}[\xi]) < 0$ because biases cause large underconsumption and/or externalities are positive, then nudges increase welfare more when $\frac{\partial x_i}{\partial n} > 0$.

where $\bar{\epsilon} = (p/\bar{x})\mathbb{E}[D_i']$ is the average price elasticity and $\mathbb{E}_W[b]$ denotes the expectation under the demand-slope-weighted distribution W, which assigns consumer i weight $w_i \equiv D_i'/\mathbb{E}[D_i']$. Thus,

$$\mathbb{E}_W[b] \equiv \mathbb{E}[w_i b_i] = \frac{\mathbb{E}[D_i' b_i]}{\mathbb{E}[D_i']},$$

so consumers whose demand is more responsive to price receive greater weight. Inverting (2) yields the central identification result:

$$\mathbb{E}_W[b] \;=\; \frac{\bar{T} \cdot p}{\bar{\epsilon} \cdot \theta} = \frac{\text{EPC} \times p}{\theta}. \tag{3}$$

Equation (3) says that under full de-biasing ($\theta = 1$), the demand-slope-weighted average bias equals the EPC times the consumer price. The equation also makes the statistic easy to generalize as we can easily vary θ to evaluate robustness to the assumption that nudges de-bias choices (see Appendix D).

We do not observe (T_i, ϵ_i) at the individual level, nor do we attempt to recover individual biases b_i. What our meta-analysis identifies is a population-weighted average of biases, where the weights are precisely the demand slopes that enter the welfare formulas. Both the treatment-effect aggregation in the data and the welfare aggregation in the planner's problem weight individuals by their marginal consumption response, so the moment our data delivers is the moment the welfare formulas require.

Distributional assumptions. Our framework derives empirical moments that can in principle be identified without any distributional assumptions on the data-generating process. In practice, however, we have found such a non-parametric approach is too demanding on the data relative to our sample sizes. Consequently, we opt for a parametric approach. We assume that $T_i > 0$ and $\epsilon_i > 0$ are jointly log-normal across consumers.[13] Write $\ln T_i \sim \mathcal{N}(\mu_T, V_T)$ and $\ln \epsilon_i \sim \mathcal{N}(\mu_E, V_E)$, with cross-correlation $\rho = \mathrm{Corr}(\ln T_i, \ln \epsilon_i)$. From (3), the log-bias inherits a normal distribution:

$$\ln b_i = \ln T_i - \ln \epsilon_i + \ln(p/\theta), \quad \mu_B = \mu_T - \mu_E + \ln(p/\theta), \quad V_B = V_T + V_E - 2\rho\sqrt{V_T V_E}.$$

Similarly, the demand slope $S_i \equiv |D_i'| = \epsilon_i \bar{x}/p$ is log-normal with $\mu_S = \mu_E + \ln(\bar{x}/p)$ and $V_S = V_E$, and the covariance between log-bias and log-slope is

$$C_{BS} = \mathrm{Cov}(\ln b_i, \ln S_i) = \rho\sqrt{V_T V_E} - V_E.$$

Under joint log-normality, all welfare-relevant cross-moments reduce to exponential functions of the five parameters $(\mu_T, V_T, \mu_E, V_E, \rho)$ plus market-level constants $(p, \bar{x}, \theta)$ and the externality distribution (μ_ξ, V_ξ). We treat ξ as random in the policymaker's information set, with prior mean $\mathbb{E}[\xi]$ and variance $\mathrm{Var}(\xi)$, reflecting epistemic uncertainty about external damages such as the social cost of carbon. The welfare-maximizing tax is $t^* = \mathbb{E}_W[b] + \mathbb{E}[\xi]$ and internalizes the expected (weighted) bias plus the expected marginal externality. Three objects enter the MVPF and excess-burden formulas:

$$\mathbb{E}[b_i S_i] = \exp\left(\mu_B + \mu_S + \tfrac{V_B + V_S + 2C_{BS}}{2}\right) = \frac{\bar{x}}{\theta}\,\mathbb{E}[T_i], \tag{4}$$

$$\mathbb{E}[b_i^2 S_i] = \exp\left(2\mu_B + \mu_S + \tfrac{4V_B + V_S + 4C_{BS}}{2}\right), \tag{5}$$

$$\mathbb{E}[S_i] = \exp\left(\mu_S + \tfrac{V_S}{2}\right) = \frac{\bar{x}}{p}\,\mathbb{E}[\epsilon_i]. \tag{6}$$

These cross-moments identify all relevant terms in the MVPF numerators:

$$\mathbb{E}_W[b] = \frac{\mathbb{E}[b_i S_i]}{\mathbb{E}[S_i]}, \qquad \mathbb{E}_W[b^2] = \frac{\mathbb{E}[b_i^2 S_i]}{\mathbb{E}[S_i]}, \qquad \mathrm{Var}_W(b) = \mathbb{E}_W[b^2] - \mathbb{E}_W[b]^2.$$

[13] For notational convenience, we work with the absolute values of treatment effects and elasticities; the signs are handled by the welfare formulas.

3.3 From data to parameters

Our meta-analysis provides study-level point estimates $(\hat{T}_j, \hat{\sigma}_{T,j})$ and $(\hat{\epsilon}_k, \hat{\sigma}_{\epsilon,k})$, where j indexes nudge studies and k indexes price studies. The observed cross-study variance of $\hat{T}_j$ conflates true heterogeneity in T_i with sampling noise. We separate the two using the law of total variance:

$$\mathrm{Var}(\hat{T}_j) = \mathrm{Var}(T_i) + \mathbb{E}[\hat{\sigma}^2_{T,j}],$$

so the noise-corrected variance is $\widehat{\mathrm{Var}}(T) = \mathrm{Var}(\hat{T}_j) - \overline{\hat{\sigma}^2_T}$, and similarly for elasticities. We then fit log-normal parameters by the method of moments: (μ_T, V_T) solve $\mathbb{E}[T] = \exp(\mu_T + V_T/2) = \overline{\hat{T}}$ and $\mathrm{Var}(T) = (e^{V_T} - 1)\exp(2\mu_T + V_T) = \widehat{\mathrm{Var}}(T)$, and likewise for (μ_E, V_E). The externality parameters (μ_ξ, V_ξ) are calibrated from external estimates, as discussed in Appendix C and reported in Table 1.

The remaining free parameter is $\rho = \mathrm{Corr}(\ln T_i, \ln \epsilon_i)$, the log-space correlation between treatment effects and price elasticities. This is the object that enters the lognormal formulas for V_B and C_{BS} derived above. From section 2.5 we know the level-space correlation $\hat{r} = \mathrm{Corr}(T_i, \epsilon_i)$. Under joint log-normality, the two are linked by

$$\hat{r} \;=\; \frac{\exp(\rho\sqrt{V_T V_E}) - 1}{\sqrt{(e^{V_T} - 1)(e^{V_E} - 1)}}, \tag{7}$$

which can be inverted to recover ρ from the empirical $\hat{r}$ and the already-estimated marginal variances V_T and V_E. We use this mapping in our baseline calibration. For markets where Section 2.5 provides no direct estimate, we assign $\hat{r}$ based on domain similarity.[14]

Market-level parameters. Finally, the welfare formulas require prices and baseline consumption levels. Table 1 in the Appendix collects these inputs, which we take from

[14] Specifically, our markets fall into two domains: health (cigarettes, alcohol, vaccines) and resources (electricity and water). We apply the cigarette correlation to the remaining health markets (alcohol and vaccines) and the electricity correlation to water. Our main results are not sensitive to this approach.

public sources for U.S. markets. To isolate the conceptual value of each policy independent of the current policy environment, we estimate all welfare effects relative to a *laissez-faire* baseline with no pre-existing taxes or subsidies. We then show how the results change once current U.S. policies are taken as the baseline.

3.4 Results

Figure 5 reports the MVPF estimates for our baseline calibration, where nudges are interpreted as de-biasing interventions ($\theta = 1$). We evaluate every policy as a departure from a laissez-faire benchmark with no pre-existing taxes, so the MVPF measures the welfare consequence of the first dollar of intervention. Panel A reports spending policies. For these policies, higher values are more favorable: a MVPF above one means that each dollar of public spending generates more than one dollar of welfare. Panel B reports revenue-raising policies. For these policies, a MVPF below one means that the government can raise one dollar while imposing less than one dollar in welfare losses on affected individuals.

Budget-Neutral Reforms We start by using the different policy tools without changing the government budget. Following Hendren and Sprung-Keyser (2020), a budget-neutral reform that finances spending policy 1 by raising revenue from policy 2 increases social welfare if and only if $\bar{\eta}_1 \mathrm{MVPF}_1 > \bar{\eta}_2 \mathrm{MVPF}_2$, where $\bar{\eta}_j$ is the average social marginal welfare weight on the beneficiaries of policy j. Within a market, the population that pays the tax and the population that receives the nudge are the same: smokers pay cigarette taxes and receive cigarette nudges, energy consumers pay energy taxes and receive energy nudges. The welfare weights cancel and the comparison reduces to $\mathrm{MVPF}_{\mathrm{nudge}} > \mathrm{MVPF}_{\mathrm{tax}}$. Under this comparison, a budget-neutral reform that raises taxes to finance more nudging is welfare-improving in energy ($2.03 > 0.34$), cigarettes

$(19.95 > 0.82)$, water $(2.33 > 0.87)$, and alcohol $(4.38 > 0.52)$, regardless of the planner's distributional preferences. For flu vaccines, where both instruments are spending policies, the analogous reallocation moves resources from the subsidy to the nudge: $\text{MVPF}_{\text{nudge}} = 62.14$ exceeds $\text{MVPF}_{\text{subsidy}} = 2.59$ by more than an order of magnitude, meaning moving budget from subsidies to nudges increases welfare.[15]

Nudges versus Taxes. To answer the question whether nudges can be more efficient than taxes, we impose one additional normative closure in Figure 6. In Panel A we assume a utilitarian welfare function so that public revenue is valued dollar-for-dollar with private surplus. Under this assumption, a dollar raised through a tax creates a welfare gain of $1 - \text{MVPF}_{\text{tax}}$, while a dollar spent on a nudge or subsidy creates a welfare gain of $\text{MVPF}_{\text{nudge}} - 1$.

In Panel B, we extend the analysis to *total* welfare effects from *non-marginal* policy changes. We maintain the utilitarian welfare function with equal welfare weights and compute the discrete welfare change of moving from a laissez-faire benchmark with no taxes and no nudges to one that either imposes the nudge or sets the tax to its optimal level (see Appendix E for details). The optimal tax equals $t^* = \mathbb{E}_W[b] + \mathbb{E}[\xi]$, which internalizes the elasticity-weighted average bias plus the expected externality.

Starting with Panel A, nudges dominate taxes at the margin in various markets. Spending a dollar on a flu vaccine nudge increases utilitarian welfare by \$61.14 per person $(p < 0.01)$, while a small increase in the subsidy on vaccines generates only 3% of that

[15] Across markets, the comparison requires distributional assumptions because payers and recipients differ. Consider raising cigarette taxes to finance flu vaccine nudges. The reform improves welfare if and only if $\bar{\eta}_{\text{flu}}/\bar{\eta}_{\text{smokers}} > \text{MVPF}_{\text{cig tax}}/\text{MVPF}_{\text{flu nudge}} = 0.82/62.14 \approx 0.013$. The threshold is two orders of magnitude below one because the flu shot nudge MVPF is so large relative to the cigarette tax MVPF: a planner who values flu vaccine recipients at any level above 1.3 percent of the value placed on smokers prefers the reform. The same logic applied to a reform that raises energy taxes to finance flu shot nudges yields a threshold of $0.34/62.14 \approx 0.005$. Across the markets we study, cross-market reforms that finance nudges out of corrective tax revenue impose weak distributional requirements: even highly unequal welfare weights would not reverse the welfare ranking.

effect, around \$1.59 in gains ($p < 0.01$). Investing another dollar on cigarette nudges generates \$18.95 of welfare per smoker, while raising another dollar through cigarette taxes generates \$0.18. The same pattern holds for alcohol, where spending a dollar on nudges raises welfare by \$3.38 per consumer, against \$0.48 from the tax.

Similarly, for resource conservation, both electricity and water nudges dominate taxes at the margin. Investing in electricity conservation nudges raises welfare by \$1 per household, while raising an additional dollar in taxes by \$0.66. These effects are statistically significantly different from zero, but as with many structural exercises confidence intervals are too wide to exclude that the difference between the two effects is significant. A similar result holds for water conservation nudges that induce \$1.33 in welfare versus only \$0.13 for taxes on the price of water.

Comparing the results of Panels A and B, the ranking between nudges and taxes can flip once we consider non-marginal effects. The two panels disagree about the policy ranking in four of five markets. In electricity, the MVPF panel ranks the nudge above the tax, but the optimal tax generates roughly *seven times* the total welfare of the optimal nudge (\$744 vs. \$104 per household per year). Flu vaccines display the same pattern: although the flu shot nudge is remarkably cost-effective and has an enormous MVPF, its total welfare effect (\$28 per person) is smaller than that of the subsidy (\$61), because the subsidy generates larger demand increases at scale. The markets for water and alcohol also see a flip in the rankings, with the nudge dominating per dollar but the tax delivering larger total surplus (\$26 versus \$23 per household per year for water, and \$278 versus \$196 per drinker per year for alcohol).

The panels agree only for cigarettes, where the nudge dominates the tax on both measures: an MVPF of 19.95 versus 0.82 at the margin, and \$104 versus \$69 per smoker per year in total surplus.

The divergence between Panel A and B is well-explained by the scale of each policy.

The MVPF is correctly identifying that each marginal nudge dollar is highly productive; the excess burden illustrates that the corrective tax operates over a vastly larger fiscal base than the nudge, which yields higher total welfare.

Policy Mix Combined policies that deploy both instruments weakly dominate either alone in every market, but Figure 6 highlights that the incremental benefit of adding the second instrument is often vanishingly small. This matters in practice. Policymakers operate under fiscal, administrative, and political constraints that often force them to deploy one instrument rather than two. *In every market we study, identifying the best instrument for the market gets society close to the first-best, and layering on a second instrument adds little.* Read in this light, our welfare estimates also speak to the concern raised by Connolly, Loewenstein and Chater (2025) that nudges carry large opportunity costs when they are used as replacements for price interventions. The pattern in our data supports that worry: in four of five markets, price instruments deliver substantially larger total surplus than nudges, and our exercise can be read as a measure of the shadow cost of the political economy constraints that favor nudges over taxes in those markets. Nudges are, at the margin, powerful additions to traditional policy, but rarely eliminate the need for price interventions.

Robustness to Alternative Models of Nudging The baseline calibration assumes nudges de-bias choices ($\theta = 1$). Meta-analytic data, as well as any observational or field data, cannot pin down the true value of θ because the same treatment effect is consistent with a small bias that the nudge fully corrects and a larger bias it only partially corrects. We therefore re-run the welfare exercise across the full partial-debiasing range $\theta \in (0, 1]$ and evaluate it at the values that Allcott, Cohen, Morrison and Taubinsky (2025) identify in a tightly controlled survey experiment that decouples nudge treatment effects from measures of choice quality. Based on their estimates, we use $\theta \approx 0.44$ for the resource markets (energy, water) and $\theta \approx 0.17$ for the health markets (cigarettes, flu vaccines, alcohol). While this changes some of our results quantitatively, the qualitative ranking is preserved in four of five markets: price instruments continue to yield larger total surplus gains than nudges in the markets for energy, residential water, influenza vaccines, and alcohol. In the cigarette market, results flip relative to our earlier finding and even here taxes start dominating nudges at scale. Thus, our overall conclusion that taxes yield larger total welfare gains than nudges is further strengthened by these robustness exercises.

Accounting for Current US Policy Our welfare analysis evaluates each instrument from a laissez-faire benchmark with no pre-existing taxes or subsidies. This isolates the conceptual value of each tool and keeps markets comparable, but it is not the right object for advising policy where corrective measures are already in place. Appendix F re-estimates the same objects taking current U.S. policy as the status quo. The two coincide in every market except where existing policy is already substantial. Cigarettes are the clear case: the current combined tax of $3.01 per pack in the US is essentially the calibrated optimum of $3.00, so all of the available corrective gain has already been realized, and an additional cigarette nudge now lowers welfare because it drives consumption below the optimum and erodes existing tax revenue. For flu vaccines the gain is only partly realized, since free vaccination captures part of an optimal subsidy that remains far higher. For energy, water, and alcohol the United States prices the relevant externalities so weakly that current policy and laissez-faire deliver nearly identical welfare, leaving the ranking between nudges and taxes unchanged.

3.5 Comparison to Other Policies in the US

A particularly useful feature of the MVPF is that it lets us compare the interventions in our data to a broader set of traditional policies in domains such as education, social insurance, and labor. Figure 7 places the spending-side MVPFs of our policies alongside category-average MVPFs from Hendren and Sprung-Keyser (2020). Their MVPFs evaluate each policy as a reform from the existing fiscal environment, so for a like-for-like comparison we evaluate ours the same way, taking current U.S. taxes and subsidies as given rather than starting from laissez-faire. Appendix F documents the current policy we take as input in each market. The gray squares report the Hendren and Sprung-Keyser category averages; the orange circles and blue triangle report our nudges and subsidies.

Evaluated at current policy, the energy and water nudges, with MVPFs of 1.95 and 2.33, remain above the category averages for adult health insurance, housing vouchers, cash transfers and tax credits, unemployment insurance, and job training, all of which fall below one. The flu shot nudge is the largest among our spending policies at 5.57, and the flu subsidy at 3.89 also exceeds every adult-policy benchmark, as well as cutting top income tax rates. The cigarette nudge is more modest at 1.53 because current cigarette taxes in the US are high and close to optimal. With taxes near the optimum, a further reduction in smoking erodes inframarginal tax revenue and lowers the nudge's value per public dollar. Alcohol nudges, at 3.41, also clear every adult-policy benchmark.

The broader comparison therefore reinforces the main lesson from Figure 5: behaviorally motivated instruments can have attractive MVPFs in the range of, or even above, important social policy tools.

4 Conclusion

What can hundreds of reduced-form studies teach us about the social value of behavioral policy? Our answer is that treatment effects become economically interpretable only when benchmarked against the fiscal instruments available in the same market. Once translated into equivalent price changes, many nudges are large: they reproduce the behavioral effects of sizable taxes or subsidies in markets for electricity, residential water, influenza vaccination, cigarettes, and alcohol consumption.

The welfare implications are more nuanced. Nudges are often highly productive per dollar of public spending and can achieve MVPFs comparable to, or above, many conventional social programs. This makes them attractive tools for marginal budget allocation and budget-neutral reforms. Yet high welfare returns per dollar do not necessarily imply large total surplus gains. Because taxes and subsidies operate at larger scale, price instruments often generate larger total surplus than nudges, even when they are less cost-effective.

The broader lesson is that behavioral interventions are neither trivial add-ons nor general replacements for conventional regulation. They are powerful instruments at the margin, but they rarely eliminate the need for price-based policy. Future research should therefore focus less on whether nudges work on average and more on the welfare primitives that determine when they should be used: who is moved by nudges, how those consumers respond to prices, whether nudges correct underlying mistakes, and how behavioral and fiscal instruments interact when deployed jointly at scale. Embedding these distinctions in large-scale program evaluations is the next step toward a more complete welfare-based science of behavioral public policy.

References

Allcott, H., Cohen, D., Morrison, W. and Taubinsky, D. (2025), 'When do nudges increase welfare?', *American Economic Review* **115**(5), 1555–1596.

Allcott, H. and Kessler, J. B. (2019), 'The welfare effects of nudges: A case study of energy use social comparisons', *American Economic Journal: Applied Economics* **11**(1), 236–76.

Alt, M., Bruns, H. and Della Valle, N. (2024), 'The more the better? synergies of prosocial interventions and effects on behavioural spillovers', *Journal of Environmental Economics and Management* **128**, 103061.

Andor, M. A., Gerster, A., Peters, J. and Schmidt, C. M. (2020), 'Social norms and energy conservation beyond the us', *Journal of Environmental Economics and Management* **103**, 102351.

Andrews, I. and Kasy, M. (2019), 'Identification of and correction for publication bias', *American Economic Review* **109**(8), 2766–2794.

Antinyan, A. and Asatryan, Z. (2019), Nudging for tax compliance: A meta-analysis, ZEW-Centre for European Economic Research Discussion Paper 19-055.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J. and Galing, S. (2017), 'Should governments invest more in nudging?', *Psychological science* **28**(8), 1041–1055.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E. and Zinman, J. (2010), 'What's advertising content worth? Evidence from a consumer credit marketing field experiment', *Quarterly Journal of Economics* **125**(1), 263–306.

Beshears, J., Choi, J. J., Harris, C., Laibson, D., Madrian, B. C. and Sakong, J. (2020), 'Which early withdrawal penalty attracts the most deposits to a commitment savings account?', *Journal of public economics* **183**, 104144.

Campos-Mercade, P., Meier, A. N., Schneider, F. H., Meier, S., Pope, D. and Wengström, E. (2021), 'Monetary incentives increase covid-19 vaccinations', *Science* **374**(6569), 879–882.

Chetty, R., Looney, A. and Kroft, K. (2009), 'Salience and Taxation: Theory and Evidence', *The American Economic Review* **99**(4), 1145–1177.

Connolly, D. J., Loewenstein, G. and Chater, N. (2025), 'An s-frame agenda for behavioral public policy research', *Behavioural Public Policy* **9**(3), 593–613.

Darby, S. et al. (2006), ‘The effectiveness of feedback on energy consumption’, *A Review for DEFRA of the Literature on Metering, Billing and direct Displays* **486**(2006).

DellaVigna, S. and Linos, E. (2022), ‘Rcts to scale: Comprehensive evidence from two nudge units’, *Econometrica* **90**(1), 81–116.

Delmas, M. A., Fischlein, M. and Asensio, O. I. (2013), ‘Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012’, *Energy Policy* **61**, 729–739.

Dubois, P., Griffith, R. and O’Connell, M. (2018), ‘The effects of banning advertising in junk food markets’, *The Review of Economic Studies* **85**(1), 396–436.

Egger, M., Davey Smith, G., Schneider, M. and Minder, C. (1997), ‘Bias in meta-analysis detected by a simple, graphical test’, *BMJ* **315**(7109), 629–634.

Farhi, E. and Gabaix, X. (2020), ‘Optimal taxation with behavioral agents’, *American Economic Review* **110**(1), 298–336.

Ferraro, P. J. and Price, M. K. (2013), ‘Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment.’, *Review of Economics and Statistics* **95**(1), 64–73.

Finkelstein, A. and Hendren, N. (2020), ‘Welfare analysis meets causal inference’, *Journal of Economic Perspectives* **34**(4), 146–167.

Gallet, C. A. and List, J. A. (2003), ‘Cigarette demand: a meta-analysis of elasticities’, *Health Economics* **12**(10), 821–835.

Grafton, R. Q., Williams, J. and Jiang, Q. (2017), ‘Possible pathways and tensions in the food and water nexus’, *Earth’s Future* **5**(5), 449–462.

Gruber, J. (2001), ‘Tobacco at the crossroads: the past and future of smoking regulation in the united states’, *Journal of Economic Perspectives* **15**(2), 193–212.

Hahn, R. W., Hendren, N., Metcalfe, R. D. and Sprung-Keyser, B. (2024), A welfare analysis of policies impacting climate change, Technical report, National Bureau of Economic Research.

Hänsel, M. C., Drupp, M. A., Johansson, D. J., Nesje, F., Azar, C., Freeman, M. C., Groom, B. and Sterner, T. (2020), ‘Climate economics support for the un climate targets’, *Nature Climate Change* **10**(8), 781–789.

Hendren, N. and Sprung-Keyser, B. (2020), ‘A unified welfare analysis of government policies’, *The Quarterly Journal of Economics* **135**(3), 1209–1318.

Houde, S. (2018*a*), 'How consumers respond to product certification and the value of energy information', *The RAND Journal of Economics* **49**(2), 453–477.

Houde, S. (2018*b*), 'The incidence of coarse certification: Evidence from the energy star program', *CER-ETH–Center of Economic Research at ETH Zurich, Working Paper* **18**, 290.

Houde, S. (2022), 'Bunching with the stars: How firms respond to environmental certification', *Management Science* **68**(8), 5569–5590.

Houde, S. and Myers, E. (2021), 'Are consumers attentive to local energy costs? evidence from the appliance market', *Journal of Public Economics* **201**, 104480.

Hummel, D. and Maedche, A. (2019), 'How effective is nudging? a quantitative review on the effect sizes and limits of empirical nudging studies', *Journal of Behavioral and Experimental Economics* **80**, 47–58.

Interagency Working Group (2021), Technical support document: Social cost of carbon, methane, and nitrous oxide interim estimates under executive order 13990, Technical report, Interagency Working Group on Social Cost of Greenhouse Gases, United States Government.

Jegnie, A., Fogarty, J. and Iftekhar, M. (2021), 'Price and income elasticities for urban residential water demand: A publication bias corrected meta-analysis'.

Manning, W. G., Keeler, E. B., Newhouse, J. P., Sloss, E. M. and Wasserman, J. (1989), 'The taxes of sin: do smokers and drinkers pay their way?', *Jama* **261**(11), 1604–1609.

Martín-Martín, A., Thelwall, M., Orduna-Malea, E. and Delgado López-Cózar, E. (2021), 'Google scholar, microsoft academic, scopus, dimensions, web of science, and opencitations' coci: a multidisciplinary comparison of coverage via citations', *Scientometrics* **126**(1), 871–906.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D. and Madrian, B. C. (2011), 'Using implementation intentions prompts to enhance influenza vaccination rates', *Proceedings of the National Academy of Sciences* **108**(26), 10415–10420.

Mullainathan, S., Schwartzstein, J. and Congdon, W. J. (2012), 'A reduced-form approach to behavioral public finance', *Annual Review of Economics* **4**(1), 511–540.

Myers, E. and Souza, M. (2020), 'Social comparison nudges without monetary incentives: Evidence from home energy reports', *Journal of Environmental Economics and Management* **101**, 102315.

Nelson, J. P. (2013), 'Meta-analysis of alcohol price and income elasticities–with corrections for publication bias', *Health economics review* **3**(1), 17.

Olmstead, S. M. (2010), 'The economics of water quality', *Review of environmental economics and policy* .

Parry, I. W., West, S. E. and Laxminarayan, R. (2009), 'Fiscal and externality rationales for alcohol policies', *The BE Journal of Economic Analysis & Policy* **9**(1), 1935–1682.

Rodemeier, M. (2025), 'Buy baits and consumer sophistication: Field evidence from instant rebates', *American Economic Journal: Economic Policy* **17**(2), 30–59.

Rodemeier, M. and Löschel, A. (2025), 'Information nudges, subsidies, and crowding out of attention: Field evidence from energy efficiency investments', *Journal of the European Economic Association* **23**(4), 1397–1430.

Rodemeier, M. and Semken, C. (2026), 'Behavioral Environmental Economics', *Available at SSRN 6509910* .

Schneider, F. H., Campos-Mercade, P., Meier, S., Pope, D., Wengström, E. and Meier, A. N. (2023), 'Financial incentives for vaccination do not have negative unintended consequences', *Nature* **613**(7944), 526–533.

Sloan, F. A., Ostermann, J., Conover, C., Taylor Jr, D. H. and Picone, G. (2004), *The price of smoking*, MIT press.

Taubinsky, D. and Rees-Jones, A. (2017), 'Attention variation and welfare: Theory and evidence from a tax salience experiment', *The Review of Economic Studies* **85**(4), 2462–2496.

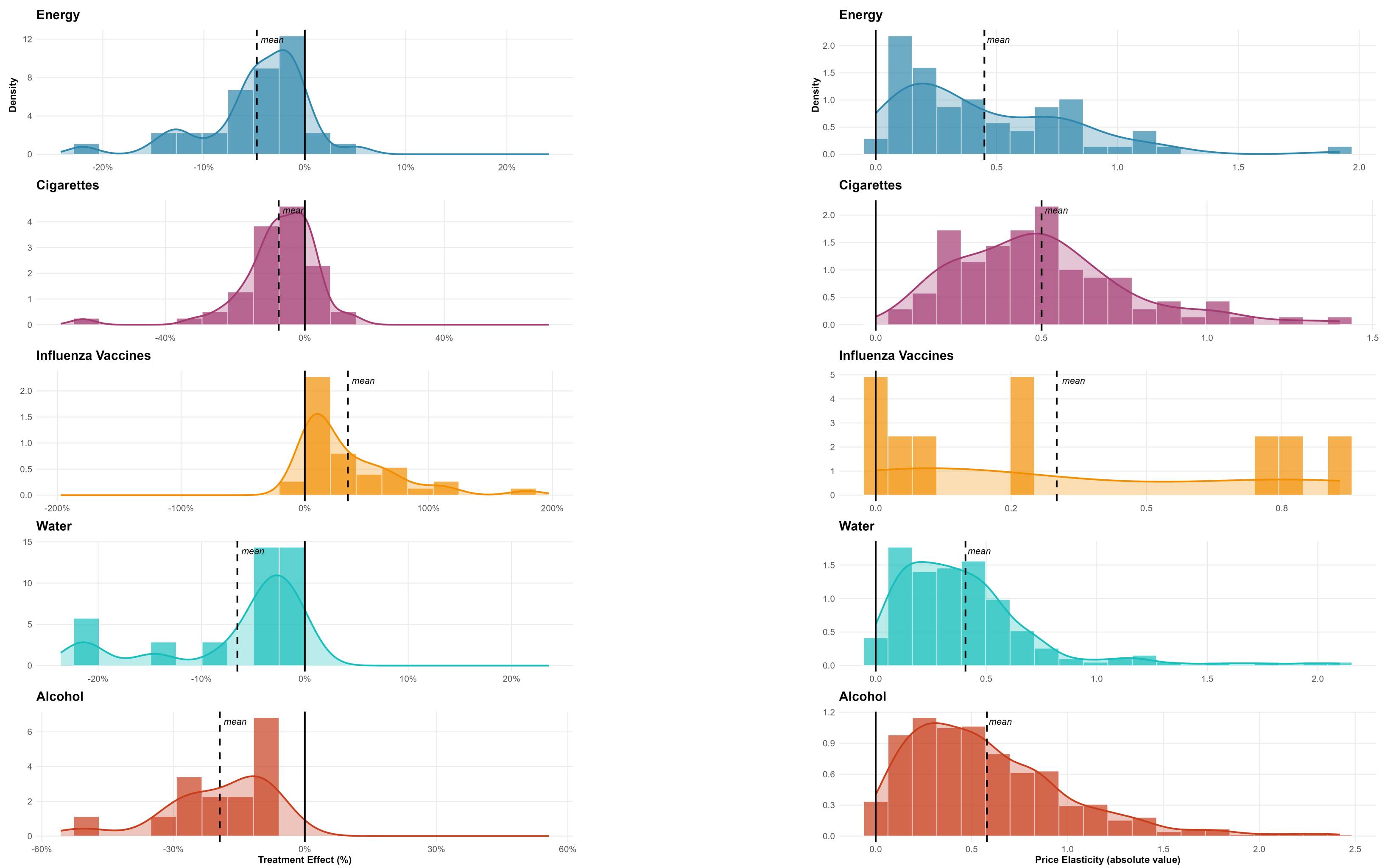
Viscusi, W. K. (1995), 'Cigarette taxation and the social consequences of smoking', *Tax policy and the economy* **9**, 51–101.

White, C. (2021), 'Measuring social and externality benefits of influenza vaccination', *Journal of Human Resources* **56**(3), 749–785.

Zhu, X., Li, L., Zhou, K., Zhang, X. and Yang, S. (2018), 'A meta-analysis on the price elasticity and income elasticity of residential electricity demand', *Journal of Cleaner Production* **201**, 169–177.

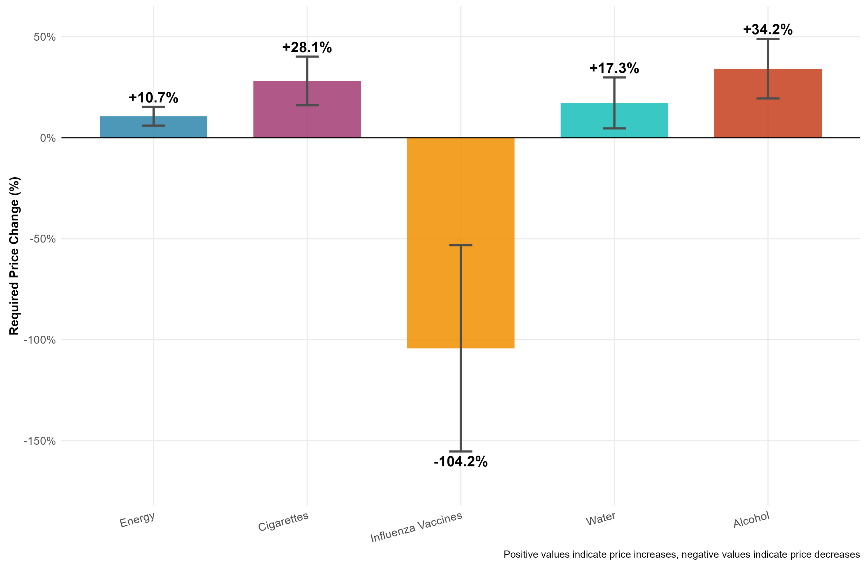
Figure 1: Distribution of Nudge Effects and Price Elasticities

(a) Nudge Effects (b) Price Elasticities



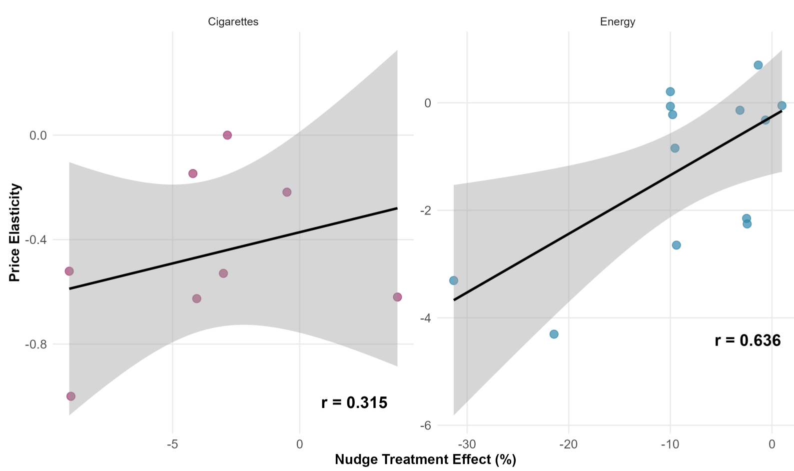
Notes: Panel (a) shows the distribution of nudge treatment effects in each market. Panel (b) shows the distribution of price elasticities (in absolute value). The solid curves are estimated log-normal distributions fitted to the data.

Figure 2: Equivalent Price Change of Nudges



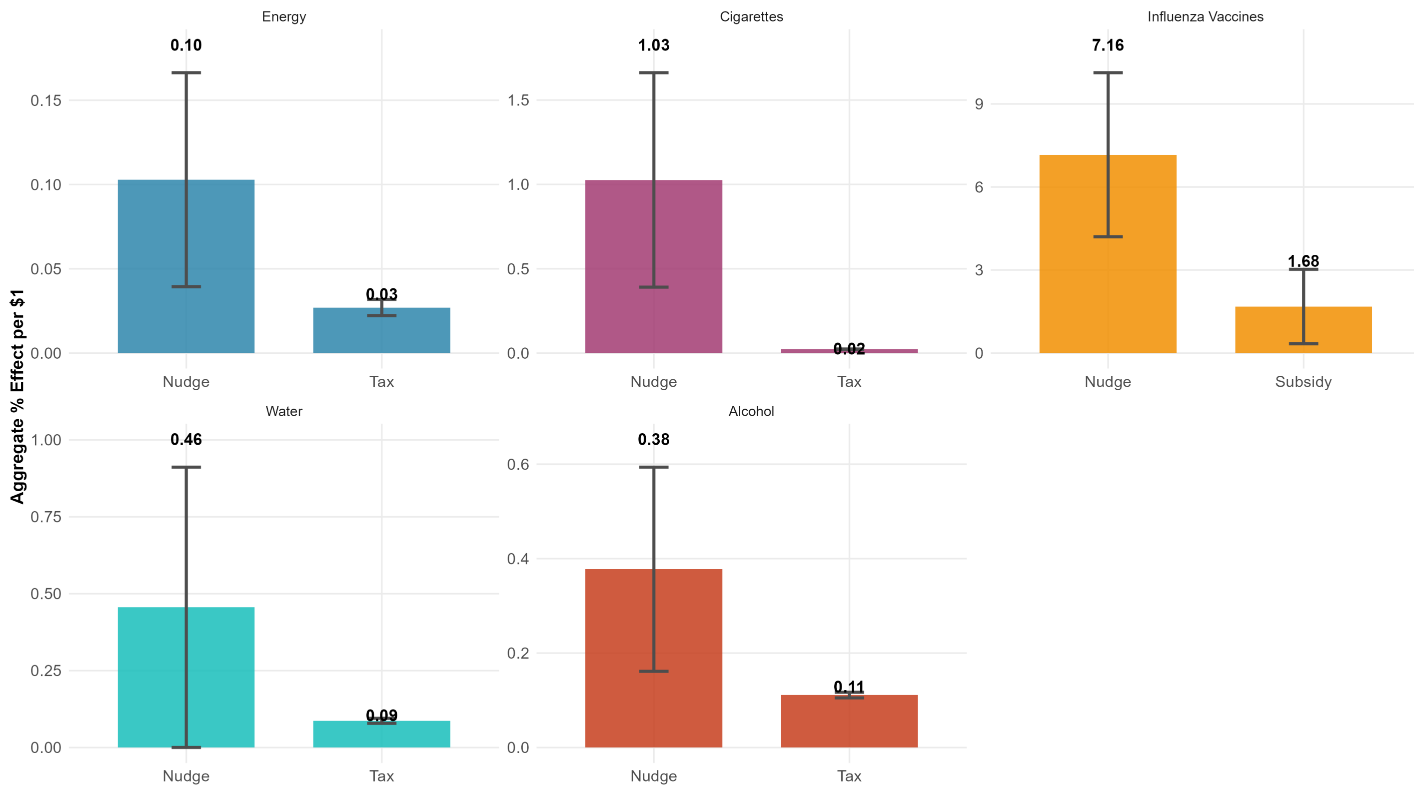
Notes: Each bar reports the equivalent price change (EPC) of the average nudge in a market, defined as the percentage change in the consumer price that would reproduce the same behavioral response as the nudge.

Figure 3: Correlation Between Price and Nudge Effects



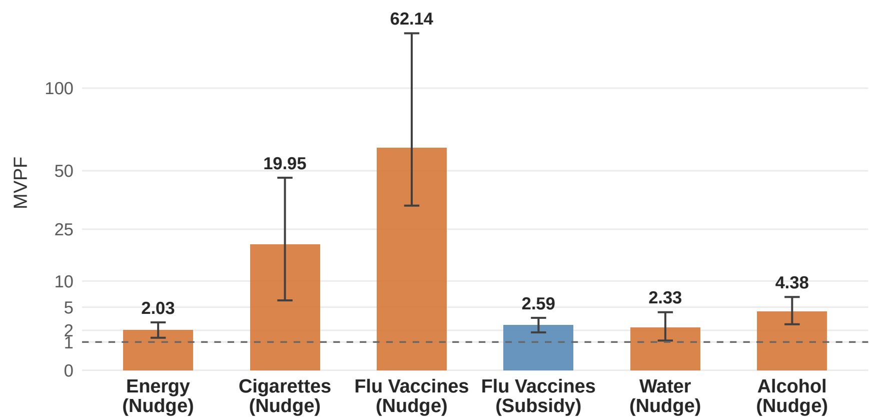
Notes: Each point represents a study that implements both a price intervention and a nudge within the same sample. The x-axis shows the nudge treatment effect (percentage change in consumption or behavior) and the y-axis shows the price elasticity.

Figure 4: Cost-Effectiveness

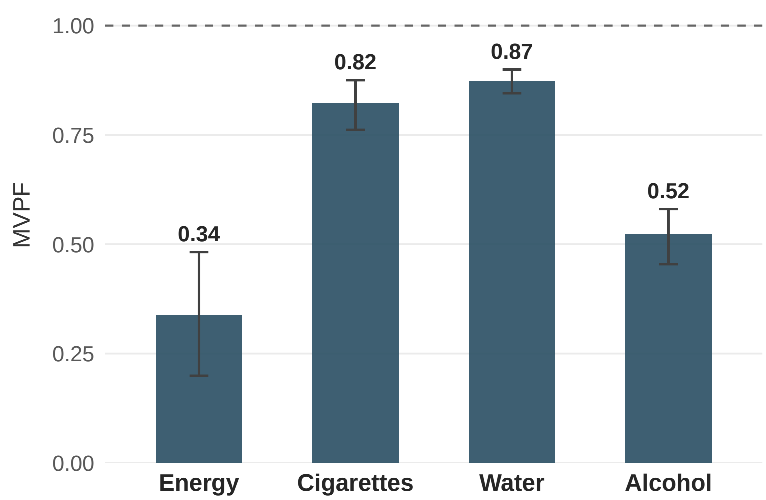


Note: The figure shows aggregate percentage behavior change per dollar invested.

Figure 5: MVPF Comparison: Revenue Raisers vs. Spending Policies



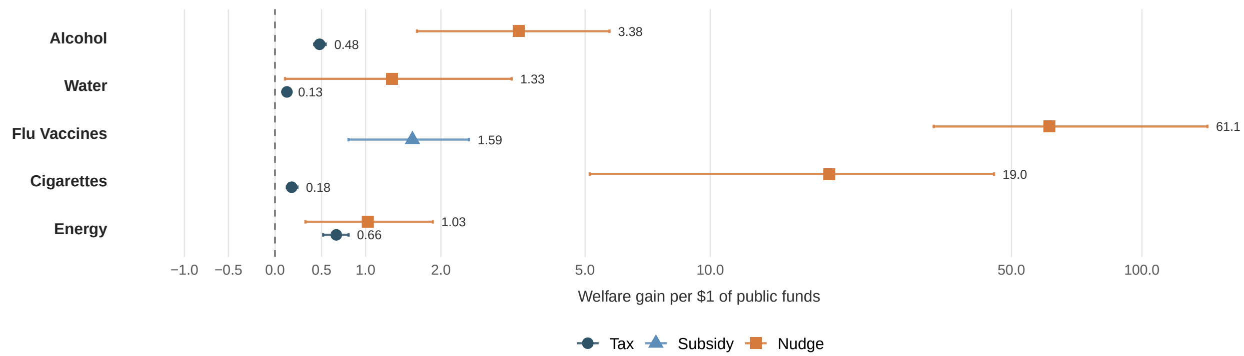
(a) Spending policies



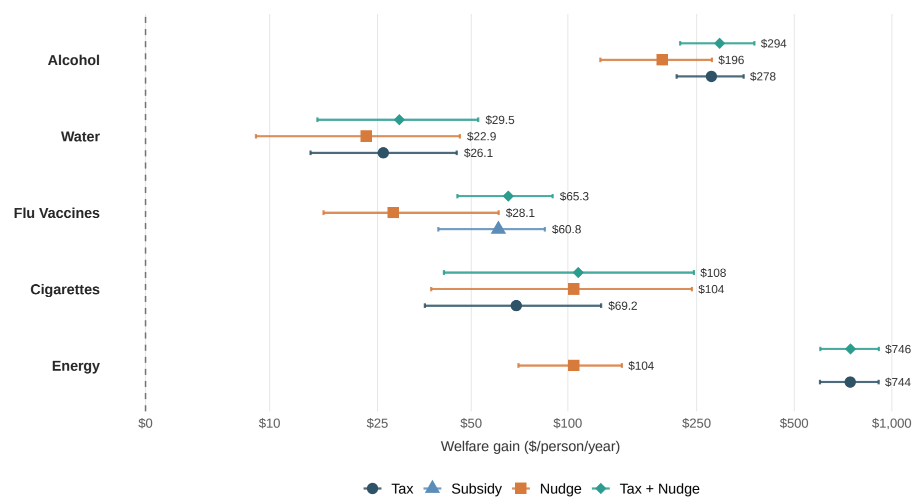
(b) Revenue raisers

Notes: Panel A reports the MVPF for nudges in all five markets, plus the flu vaccine subsidy; values above one indicate that the policy is welfare-improving. Panel A uses a square-root y-axis so that the smaller nudge MVPFs remain legible alongside the large cigarette and flu vaccine nudge MVPFs. Panel B reports the marginal value of public funds (MVPF) for a tax on energy, cigarettes, water, and alcohol; values below one indicate that a tax is welfare-improving. Whiskers are 95% percentile bootstrap confidence intervals from 2,000 replications.

Figure 6: Utilitarian Welfare Effects on a Common Dollar Scale



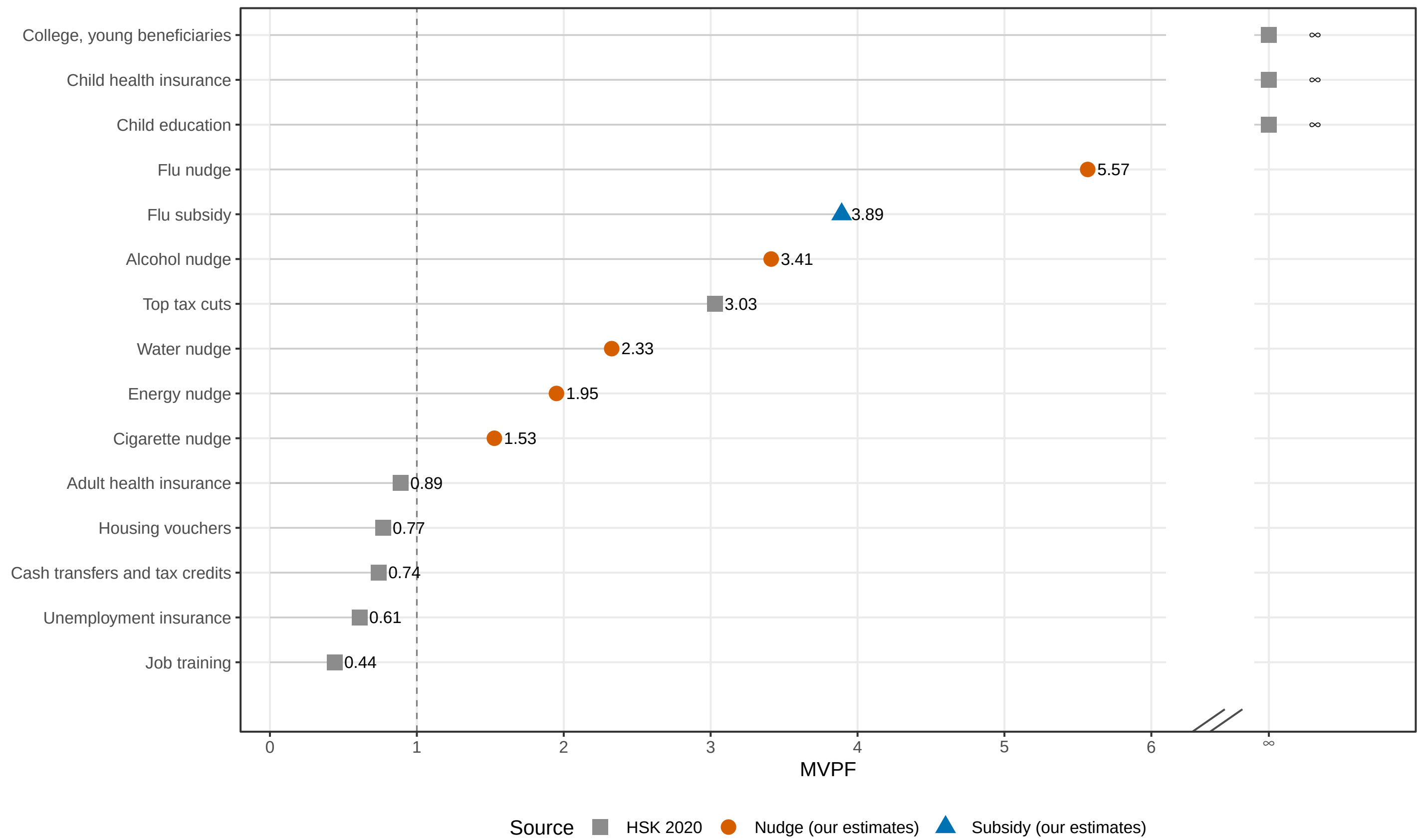
(a) Marginal Welfare Gains



(b) Total Welfare Gains

Notes: For revenue-raising taxes, Panel A plots $1 - \mathrm{MVPF}$. For spending policies, including nudges and the flu vaccine subsidy, it plots $\mathrm{MVPF} - 1$. Values above zero indicate positive marginal welfare gains on this common scale. Panel A uses a pseudo-log x-axis so that the large flu-vaccine nudge estimate can be shown together with estimates near zero and below zero. Error bars show 95% confidence intervals from 2,000 bootstrap replications.

Figure 7: Behavioral Policies in Perspective



Notes: The figure compares the spending-side MVPFs of our behavioral policies, evaluated at current U.S. policy, with category-average MVPFs from Hendren and Sprung-Keyser (2020), taken from their Table II and Figure IV, Panel B. Both are measured relative to the existing policy environment, taking current taxes and subsidies as given. Appendix F details current US policies we take as input. Orange circles denote nudges and blue triangles subsidies. Higher MVPFs indicate larger welfare gains per dollar of government cost.

ONLINE APPENDIX

A The Top Performers among Nudges

In this Appendix, we discuss which nudge types are most effective in each market. Figure 8 plots the nudge categories with the largest treatment effects in each market. For cigarette consumption, information interventions have the largest category-level effect, reducing smoking probability by 9.69%. Feedback interventions reduce smoking by 3.52%, followed by other nudges at 2.41%.

For vaccines, reminders to get vaccinated have the largest effect equal to a 115% increase in the probability to get vaccinated. Defaults—such as pre-scheduling vaccination appointments that patients must actively cancel—cause similarly large effects (105%), followed by information provision about the benefits and importance of flu shots (68.9%).

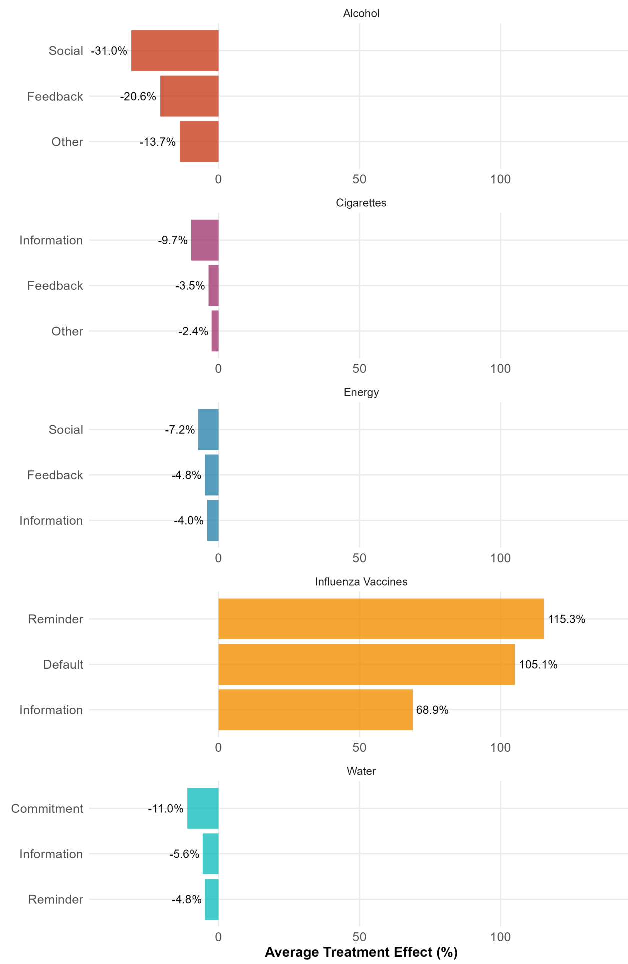
Household energy nudges are most effective when they provide social comparisons, which reduce electricity consumption by 7.17%. Feedback interventions reduce consumption by 4.84%, followed by information interventions at 4.02%.

In the water market, commitment and goal-setting interventions are the top performers, reducing consumption by 11.0% on average. Information provision reduces consumption by 5.60%, and reminders reduce consumption by 4.80%.

Alcohol consumption is most responsive to social-norm interventions (31.0%) and feedback interventions (20.6%), followed by other brief interventions (13.7%) and information interventions (11.0%).

Taken together, the most effective nudge varies systematically across markets, suggesting that different behavioral frictions may bind in different domains: inattention and biased beliefs in smoking, forgetfulness and hassle costs in vaccination, social-image or norm-based motives in energy and alcohol, and self-control and biased beliefs in resi-

Figure 8: Most Effective Nudge Category by Market



dential water conservation.

B Publication Bias

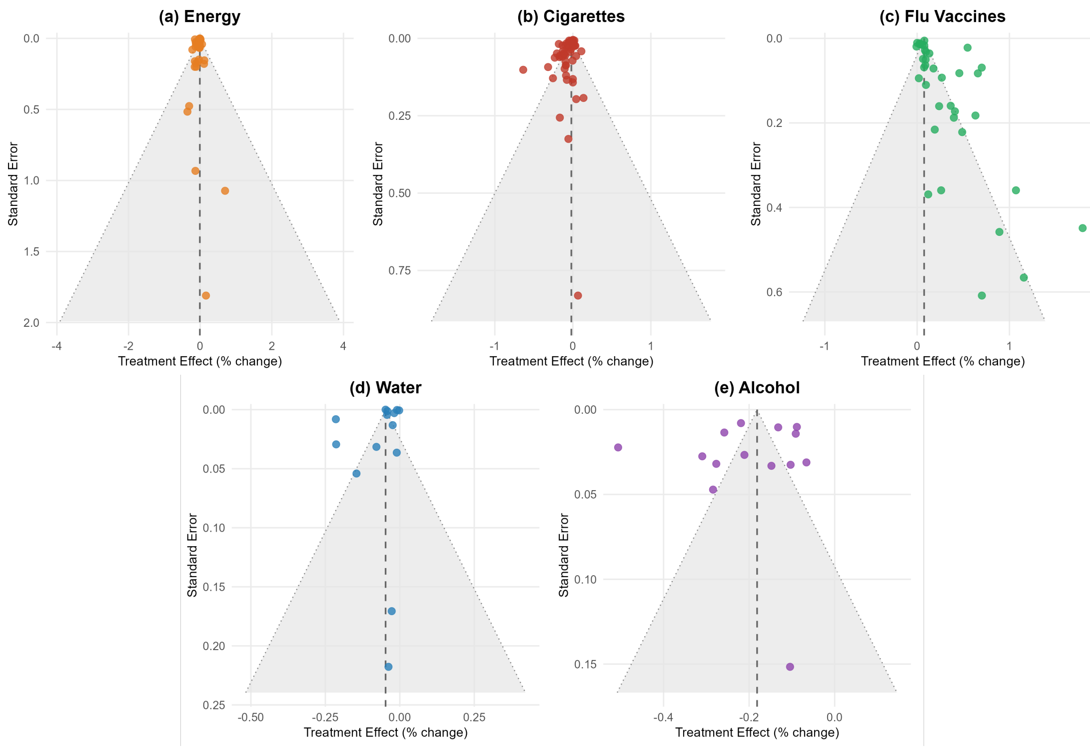
A common concern with meta-analyses is that published studies overrepresent statistically significant results, inflating pooled effect sizes. We assess publication bias using funnel plots and the Egger regression test (Egger et al. 1997). A funnel plot displays each study's treatment effect against its standard error: under no selection and a homogeneous true effect, estimates should scatter symmetrically around the mean, with imprecise studies fanning out at the bottom and precise studies clustering at the top. Egger's test formalizes this by regressing the standardized effect on precision; a nonzero intercept indicates asymmetry of the kind selective publication would produce.

Figure 9 shows funnel plots for all five markets, and Egger tests detect significant asymmetry in three: energy ($t = -2.75$, $p = 0.009$), cigarettes ($t = -3.53$, $p < 0.001$), and flu vaccines ($t = 2.46$, $p = 0.019$). Water ($t = 1.30$, $p = 0.217$) and alcohol ($t = -0.66$, $p = 0.525$) show no significant asymmetry, though small sample sizes limit power. The signs of the intercepts indicate that smaller studies tend to report larger effects in the hypothesized direction. Taken together, the evidence suggests that published nudge effects in energy, cigarettes, and flu vaccines are likely overstated relative to the true underlying effect.

We show that our main estimates of EPC, cost effectiveness, and MVPF are robust to trying to explicitly model and correct for publication bias. Andrews and Kasy (2019) recently provided a framework showing how to nonparametrically identify publication bias under the assumption that true treatment effect sizes are independent of underlying standard errors. Their full nonparametric approach places stringent requirements on the size of the underlying data, so the authors suggest a simple parametric approximation where selective publication depends only on if the main effect of the paper is statistically significant. Publication bias can then be identified by measuring the degree to which the density of estimates with t-statistics with magnitude just above 1.96 exceeds the density of estimates with t-statistics just under 1.96.[16] Let $R > 1$ denote this ratio. We can correct for publication bias by running all of our main specifications on a weighted sample, where papers with non-significant estimates receive a weight of R while papers with

[16] In practice, we compute the number of papers with t-statistics in $(1.96, 2.46)$ divided by the number of papers with t-statistics in $(1.46, 1.96)$.

Figure 9: Funnel Plots and Publication Bias Tests



Notes: Panels (a)–(e) show funnel plots for nudge treatment effects in each of the five markets. Each point represents a study-level estimate. The dashed vertical line indicates the inverse-variance weighted mean effect, and the gray shaded region shows the 95% pseudo-confidence interval under the assumption of no publication bias. Precise studies (small standard errors) appear at the top.

significant estimates receive a weight of 1.

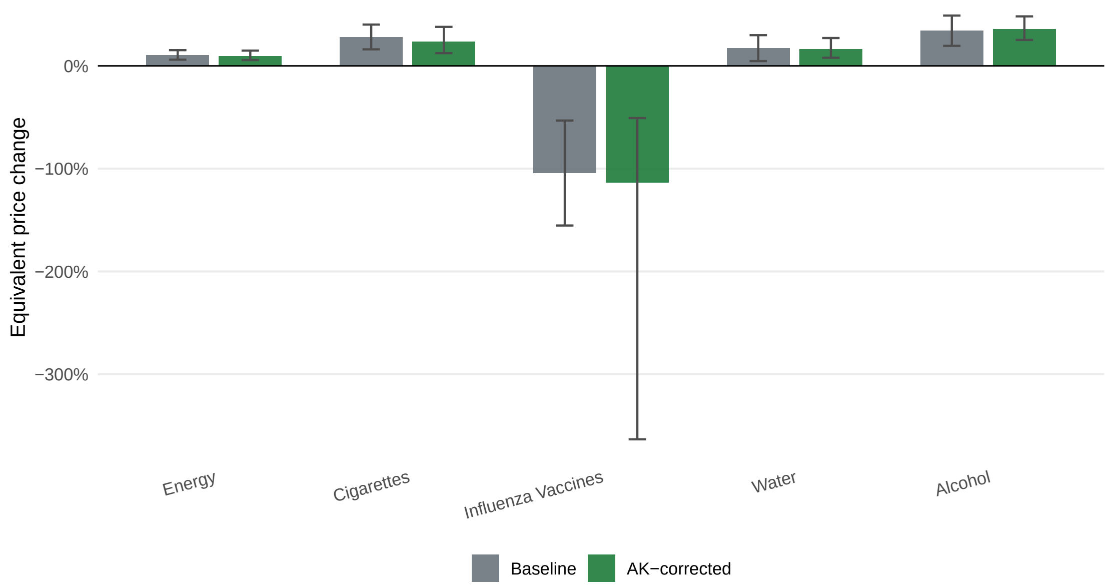
We separately estimate elasticity publication bias and nudge treatment effect publication bias. The former is estimated at $\hat{R}_{elas} = 1.59$, while nudge treatment effect publication bias is $\hat{R}_{nudge} = 1.64$. This implies that in both literatures, there is some evidence for modest publication bias, and in addition, nudge studies exhibit slightly larger publication bias. This could bias our estimates of EPC upwards, which in turn would tend to imply that our conclusions about nudges may be too optimistic.

We therefore re-estimate EPC, cost-effectiveness, and MVPF of nudges and taxes in our 5 markets, after correcting for publication bias. Figures 10—12 compare the main estimates reported in our paper against the publication bias corrected versions. We compute standard errors by constructing bootstrap redraws of our entire dataset and re-estimating publication bias in addition to the main estimates. Comparing point estimates, we find that the relative ranking between nudges and taxes remains similar when accounting for publication bias. Quantitatively, most point estimates are hardly changed when accounting for publication bias. The one exception is the cigarette market, where accounting for publication bias substantially reduces the MVPF associated with nudges. Even here, however, the MVPF still remains above the MVPF of taxes, consistent with our findings in the body of the paper.

Accounting for publication bias also substantially increases the variance of all estimates. This is because even given the parametric specification suggested by Andrews and Kasy (2019), and pooling across markets, $\hat{R}_{elas}$ and $\hat{R}_{nudge}$ remain fairly large.

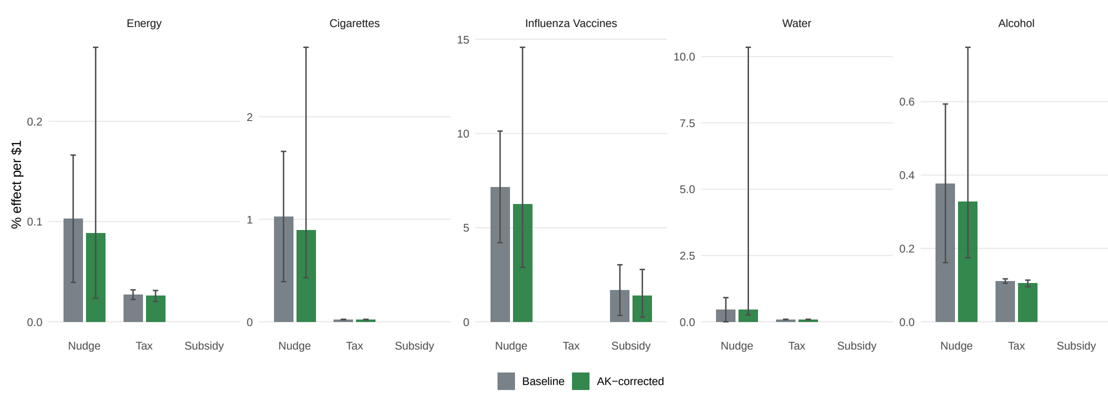
Overall, our key results and policy conclusions are robust to the correction for publication bias.

Figure 10: Publication Bias Corrections: EPC



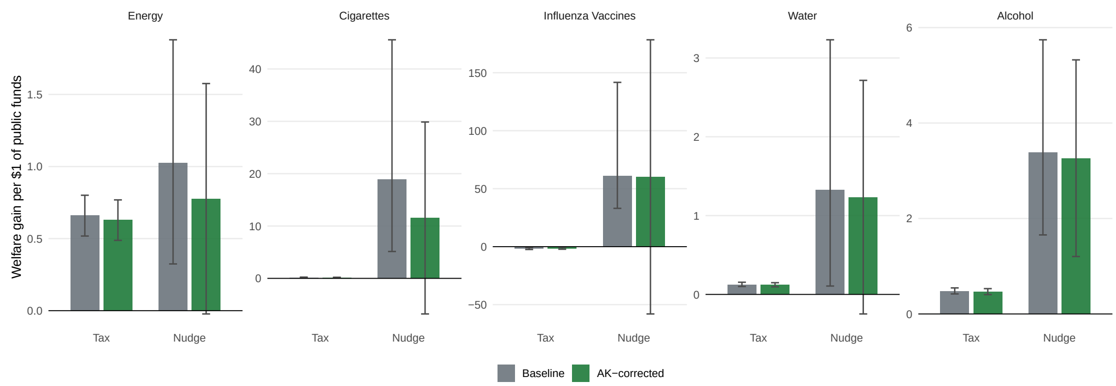
Notes: Panels (a)–(e) compare EPC estimates with and without publication bias corrections in each of the five markets. Each bar represents point estimates, while error-bars represent 95% bootstrapped confidence intervals. Bootstrap draws redraw both the data used to estimate publication bias ($\hat{R}_{elas}$ and $\hat{R}_{nudge}$) as well as data used to estimate treatment effects, given publication bias estimates.

Figure 11: Publication Bias Corrections: Cost Effectiveness



Notes: Panels (a)–(e) compare cost-effectiveness estimates for nudges and taxes, with and without publication bias corrections in each of the five markets. Each bar represents point estimates, while error-bars represent 95% bootstrapped confidence intervals. Bootstrap draws redraw both the data used to estimate publication bias ($\hat{R}_{elas}$ and $\hat{R}_{nudge}$) as well as data used to estimate treatment effects, given publication bias estimates.

Figure 12: Publication Bias Corrections: MVPF



Notes: Panels (a)–(e) compare MVPF estimates for nudges and taxes, with and without publication bias corrections in each of the five markets. Each bar represents point estimates, while error-bars represent 95% bootstrapped confidence intervals. Bootstrap draws redraw both the data used to estimate publication bias ($\hat{R}_{elas}$ and $\hat{R}_{nudge}$) as well as data used to estimate treatment effects, given publication bias estimates.

C Calibrated Market Parameters

Table 1 lists the calibrated market parameters we use as inputs for our welfare computations. For electricity consumption externalities, we use a social cost of carbon of \$181/tCO_2, taken from the study by Hänsel et al. (2020). The same data yields a standard deviation of \$186/tCO_2. As a robustness check we re-run our results at the lower \$51/tCO_2 value used by the U.S. federal government for cost–benefit analysis (Interagency Working Group 2021). The qualitative ranking between nudges and taxes is unchanged.

For cigarettes, the per-pack externality of \$0.68 averages the external-cost estimates in Sloan et al. (2004), Viscusi (1995), and Gruber (2001). The relevant components are secondhand-smoke harms imposed on non-smokers and external healthcare costs net of the offsetting "death benefit" to public pension and long-term-care systems.

Table 1: Market Parameters

	Energy	Cigarettes	Flu Vaccines	Water	Alcohol
Prices and consumption					
Price	\$0.13/kWh	\$8.00/pack	\$41/shot	\$11.50/CCF	\$1.27/10 g
Consumption level	11,000 kWh/yr	260 packs/yr	0.39 (rate)	100 CCF/yr	1,200 $\times$10 g/yr
Externalities					
Per-unit externality (ξ)	\$0.19/kWh	\$0.68/pack	\$153/shot	\$2.00/CCF	\$0.57/10 g
Externality SD	\$0.19	\$0.09	\$45	\$0.50	\$0.08
Meta-analysis estimates					
Avg. nudge effect	-4.75%	-7.49%	$+34.79\%$	-6.53%	-19.10%
Avg. price elasticity	-0.45	-0.49	-0.33	-0.38	-0.57
N (nudge estimates)	35	53	36	14	15
N (price estimates)	68	94	9	174	560
Correlation (r)	0.636	0.315	0.315	0.636	0.315
Nudge implementation costs					
Avg. cost per subject	\$53.96	\$8.89	\$0.55	\$11.90	\$57.26

Notes: Externality estimates are taken from the literature as explained in this appendix. The correlation r uses empirical paired-study correlations for energy and cigarettes; the cigarette correlation is applied to the other health markets (flu vaccines and alcohol) and the energy correlation to water. Implementation costs are averages across studies reporting costs. Consumption levels are computed per consumer rather than per capita (e.g., per smoker for cigarettes). For alcohol, quantities are expressed per 10 g of pure ethanol (the WHO standard-drink unit; a US standard drink is 14 g), measured per drinker.

Influenza vaccination has a positive externality through reduced transmission to unvaccinated individuals. White (2021) estimates the social benefit of vaccination using an instrumental-variables strategy that exploits variation in flu-shot uptake, reporting per-shot external benefits of \$63 under a conservative value of a statistical life and approximately \$240 under the EPA value. We take \$153 per shot as the mean, and treat the range as an approximate 95% interval, giving a standard deviation of $(240 - 63)/(2 \times 1.96) = 45.15 \approx \$45$.

Unlike the carbon and tobacco literatures, no single canonical estimate exists for the per-unit external cost of residential water. We aggregate four credible components: (i) the carbon and energy costs of pumping, treating, and distributing water; (ii) nutrient and contaminant pollution from wastewater discharge into watersheds; (iii) extraction rents in scarcity-constrained regions; and (iv) ecosystem-service losses associated with reduced in-stream flows. Drawing on Olmstead (2010) and Grafton, Williams and Jiang (2017), the component estimates range from approximately \$1.10 to \$2.80 per hundred cubic feet (CCF). We take the midpoint, \$2.00, as the mean, and treat the range as the support of a uniform distribution, giving a standard deviation of $(2.80 - 1.10)/\sqrt{12} = 0.49 \approx \$0.50$.

The alcohol externality captures external costs such as drunk-driving injuries and fatalities imposed on other road users, property damage, criminal-justice costs, and external medical costs. Manning et al. (1989) provide the foundational estimate of \$0.48 per fluid ounce of ethanol (1986 dollars), inflation-adjusted to \$1.34/oz in 2024 dollars. At 23.33 g per fluid ounce of pure ethanol this is \$0.057 per gram, or \$0.57 per 10 g of ethanol (about \$0.80 per US standard drink), the unit we use throughout. Parry, West and Laxminarayan (2009) provide an independent corroborating estimate of \$0.44 per 10 g using a similar external-cost methodology, with the difference driven by alternative drunk-driving fatality assumptions. The standard deviation of \$0.08 per 10 g reflects the cross-study range.

D Robustness to Alternative Models of Nudges

Our leading specification follows the reduced-form approach to behavioral public finance (Mullainathan, Schwartzstein and Congdon 2012) and uses observed treatment effects to infer the underlying bias. In our notation, this is the case $\theta = 1$ (full debiasing) and $\kappa = 0$ (no psychic cost). In reality, nudges may close only a fraction of the underlying bias, $\theta < 1$, or even exacerbate it. Meta-analytic data alone cannot pin down either parameter from the observed treatment effect, and even most experiments are unable to speak to these parameters. The one exception is the recent study by Allcott, Cohen, Morrison and Taubinsky (2025) who run two tightly controlled experiments that decouple the treatment effect of an information label from the underlying bias the label is designed to correct—given a set of auxiliary assumptions that we do not discuss here. Their methodology recovers parameters that our meta-analytic moments cannot identify: the average debiasing intensity, θ, a (self-reported) measure of κ, and finally, a measure of whether nudges are well-targeted to biased consumers. They find that in both markets the labels close only a fraction of the bias on average: $\theta \approx 0.44$ for fuel-economy labels in their cars experiment and $\theta \approx 0.17$ for sugary-drink warning labels in their drinks experiment. For κ, the estimate they use for calibrations across interventions is approximately zero. Finally, they do not find evidence that nudges target more biased consumers.

We test the robustness of our welfare results to these findings. While $\kappa = 0$ does not change our main results, we must account for the finding that nudges may only partially de-bias choices. Figure 13 reports the result. Each panel plots welfare gains under the optimal tax (or subsidy), the nudge, and the combined policy as a function of θ. For illustrative purposes, the vertical purple line marks the estimate in Allcott, Cohen, Morrison and Taubinsky (2025) for the respective market: $\theta = 0.44$ for the resource markets (energy, water) and $\theta = 0.17$ for the health markets (cigarettes, flu vaccines, alcohol).[17]

At lower θ, both the inferred bias and welfare gains for both instruments rise. To understand this pattern, recall how the behavioral bias is identified: The magnitude of the behavioral bias is increasing (linearly) in the nudge treatment effect. This implies that the bias is larger if the nudge only imperfectly de-biases consumers than if it were fully de-biasing. For example, if a nudge reduces the deviation in consumption from the private optimum by 50%, then the distortion caused by the behavioral bias is twice as large as the treatment effect. Given some treatment effect, lower values of nudge effectiveness, therefore, imply that the magnitude of the behavioral bias is larger. In the case of taxation, this mechanically implies that the gains from corrective taxation are larger for lower θ. In the case of nudging, the directional effect of a change in nudge effectiveness is ambiguous. On the one hand, lower nudge effectiveness implies a larger bias, which increases the benefits of nudging. But on the other hand, lower nudge effectiveness also implies that the nudge is less powerful at de-biasing, which decreases its benefits. Interestingly, our results suggest that the former effect dominates the latter in all markets.

The qualitative ranking of nudges versus the corresponding price instrument is preserved across the entire partial-debiasing in four out of five markets. Price instruments continue to yield larger total welfare in the markets for household energy, residential water, influenza vaccines, and alcohol. This is not only true for the θ values in Allcott, Cohen, Morrison and Taubinsky (2025) but for *all* values. The one exception where our results flip is the market for cigarettes. Here, taxes start dominating nudges if the latter is only partially de-biasing.

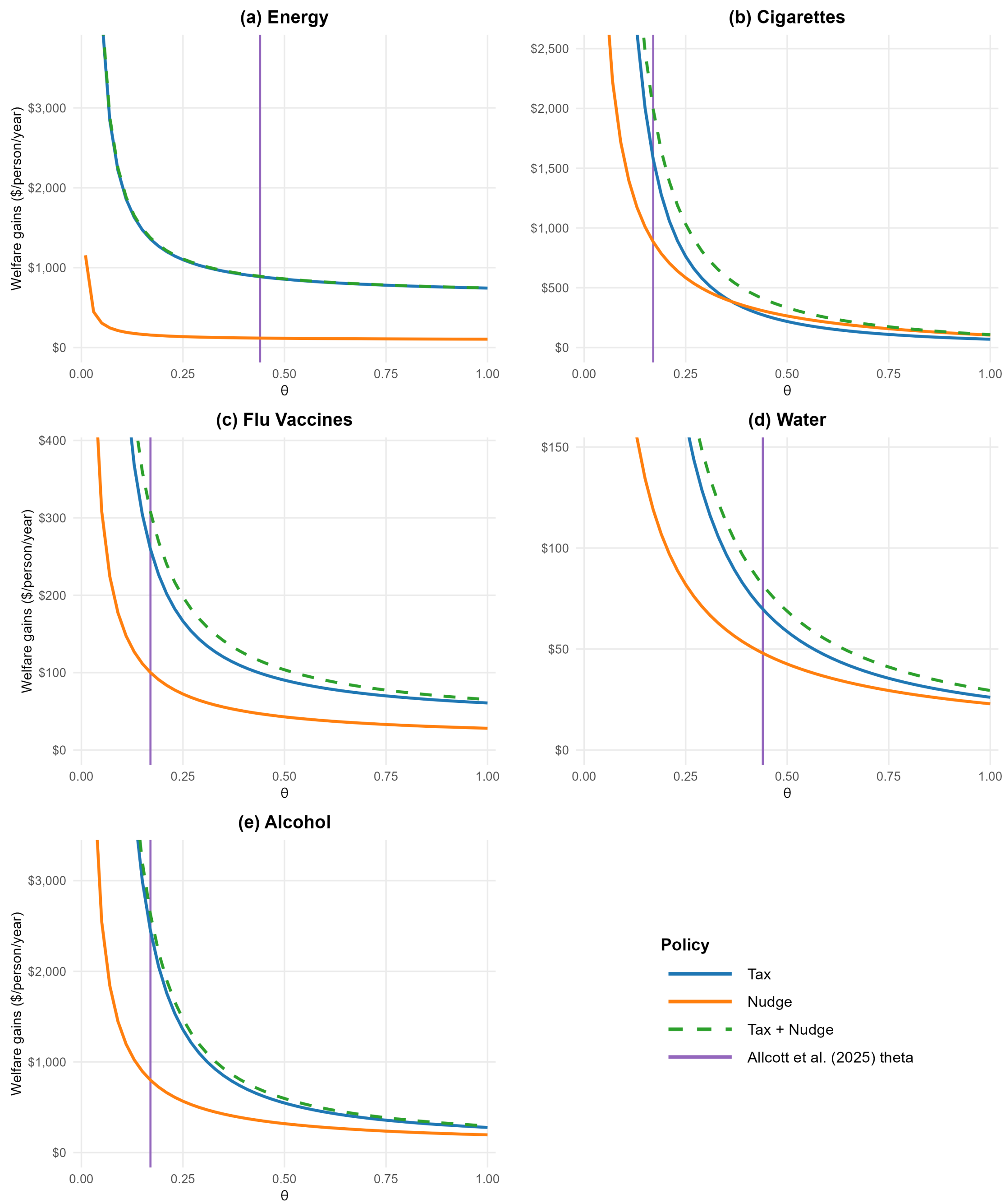
We have not yet addressed the finding in Allcott, Cohen, Morrison and Taubinsky

[17]Thus, we use their estimates from the fuel economy experiment to inform resource markets in our meta-analysis (energy, water) and their estimates from the sugar drinks market to inform health markets (cigarettes, flu vaccines, alcohol).

(2025) that nudges may increase the variance in bias. This increase in variance unambiguously reduces the welfare attributed to nudges by misallocating consumption across consumers. Two implications follow for our work. First, the welfare gains we report under partial debiasing here are an upper bound on the welfare gains a noise-adjusted calibration would deliver. The variance penalty would shift the nudge curve in Figure 13 downward at every θ, leaving the corrective tax and subsidy curves unchanged. Second, this shift can only strengthen the qualitative ranking in markets where the corrective tax already dominates the optimal nudge (energy) and can only erode the nudge's advantage in markets where the nudge dominates the corrective tax (cigarettes). In every direction the variance result moves the welfare comparison toward the price instrument and away from the nudge. Our qualitative policy conclusions survive both empirical departures from the leading specification.

Overall, allowing nudges to be only partially debiasing, or imperfectly targeted, leaves our main qualitative conclusion intact that price instruments provide larger total surplus in most markets.

Figure 13: Welfare under Partial Debiasing



Notes: Welfare gains in dollars per person or household per year as a function of the de-biasing intensity θ.

E Derivations of Utilitarian Welfare Effects

This appendix collects the formal derivations underlying the results for the total surplus effects in a utilitarian framework, reported in the main text. We restate the consumer's problem, define the welfare-weighted moments used throughout, derive the welfare function $W(t,n)$ as a quadratic loss around the first-best allocation, and then specialize the formula to the three cases discussed in Section 3: the optimal Pigouvian tax alone, the optimal de-biasing nudge alone, and the optimal combination. We close with a formal statement and proof of the nudge-dominance condition $\mathrm{sd}_W(b) \geq |\mathbb{E}[\xi]|$ and an extension to the moral-tax model.

E.1 Setup

Consider a market with a unit mass of heterogeneous consumers indexed by i, with quasi-linear preferences $u_i(x_i) + y_i$ over a consumption good x_i and a numeraire y_i. The consumer price is $p = c + t$, where c is the constant marginal cost and t the corrective tax. We assume u_i is twice continuously differentiable with $u_i'' < 0$, and we work in a neighborhood where the budget constraint binds with strict interior consumption.

Each consumer solves a perceived first-order condition shifted by a money-metric bias b_i:

$$u_i'(x_i^*(t,n)) = p + \kappa n - (1 - \theta n)\, b_i, \tag{8}$$

where $n \in [0,1]$ is the nudge parameter ($n = 0$ no nudge, $n = 1$ the deployed nudge we measure in the data), $\theta \in \mathbb{R}$ is the de-biasing intensity, and $\kappa \geq 0$ is the per-unit psychic cost imposed by the nudge. Let $D_i' \equiv \partial x_i^* / \partial p < 0$ denote consumer i's demand slope. Each unit of consumption generates a per-unit externality ξ that enters social welfare but not private choices, with $\xi \perp (b_i, D_i')$.

The first-best allocation x_i^{FB} satisfies

$$u_i'(x_i^{FB}) = c + \mathbb{E}[\xi], \tag{9}$$

i.e., marginal utility equals marginal social cost. The wedge between the privately and socially optimal first-order conditions, evaluated at $x_i^*(t, n)$, is therefore

$$\Delta_i(t,n) \;\equiv\; t + \kappa n - \big((1-\theta n)b_i + \mathbb{E}[\xi]\big). \tag{10}$$

At $\Delta_i = 0$ for all i, the consumer's marginal utility coincides with the social marginal cost and her chosen allocation equals x_i^{FB}.

Welfare-weighted moments. For any random variable Z_i defined on the population of consumers, define the demand-slope-weighted expectation

$$\mathbb{E}_W[Z] \;\equiv\; \frac{\mathbb{E}[\,|D_i'|\,Z_i\,]}{\mathbb{E}[\,|D_i'|\,]} \;=\; \mathbb{E}[w_i Z_i]\,, \qquad w_i \;\equiv\; \frac{|D_i'|}{\mathbb{E}[\,|D_i'|\,]}. \tag{11}$$

The weights $w_i \geq 0$ satisfy $\mathbb{E}[w_i] = 1$. The welfare-weighted variance and standard deviation are

$$\mathrm{Var}_W(Z) \;\equiv\; \mathbb{E}_W[Z^2] - \mathbb{E}_W[Z]^2, \qquad \mathrm{sd}_W(Z) \;\equiv\; \sqrt{\mathrm{Var}_W(Z)}.$$

These weights reflect responsiveness: consumers whose behavior is more sensitive to price contribute more to welfare, since a given bias produces a larger consumption distortion when $|D_i'|$ is large.

Quadratic-utility approximation. We adopt the standard local approximation that demand is linear, or equivalently, that u_i is locally quadratic. Concretely, in a neighborhood of x_i^{FB},

$$u_i(x_i) \;\approx\; u_i(x_i^{FB}) + u_i'(x_i^{FB})(x_i - x_i^{FB}) + \tfrac{1}{2}u_i''(x_i^{FB})(x_i - x_i^{FB})^2, \qquad (12)$$

with $u_i''(x_i^{FB}) = 1/D_i' < 0$. The derivations below should be read as exact under this quadratic environment, or equivalently as second-order Taylor approximations to welfare under any smooth utility with locally linear demand.

E.2 Derivation of the Welfare Function

We now derive the social welfare function $W(t, n)$ as a function of the policy pair (t, n). Social welfare aggregates rational consumer surplus net of psychic costs, plus tax revenue, minus externality damages:

$$W(t, n) \;=\; \mathbb{E}[\, u_i(x_i^*) - (c + \mathbb{E}[\xi])x_i^* \,] \;-\; \kappa n \,\mathbb{E}[x_i^*]. \qquad (13)$$

The term $u_i(x_i^*) - (c + \mathbb{E}[\xi])x_i^*$ is total surplus consumer i would generate at her chosen quantity if she internalized the externality, i.e., the social-planner objective evaluated at the actual allocation. The transfer terms $-px_i^* + tx_i^*$ from consumer surplus and government revenue cancel against $-cx_i^*$ leaving $-(c+\mathbb{E}[\xi])x_i^*$ as the net of consumption and externality. The final $-\kappa n\mathbb{E}[x_i^*]$ is the aggregate psychic cost, which is destroyed rather than redistributed and therefore enters welfare directly.

The first-best welfare is $W^* = \mathbb{E}\big[u_i(x_i^{FB}) - (c + \mathbb{E}[\xi])x_i^{FB}\big]$. Substituting the quadratic expansion (12) into (13), using $u_i'(x_i^{FB}) = c + \mathbb{E}[\xi]$, and writing $\delta_i \equiv x_i^* - x_i^{FB}$,

$$u_i(x_i^*) - (c + \mathbb{E}[\xi])x_i^* \;=\; u_i(x_i^{FB}) - (c + \mathbb{E}[\xi])x_i^{FB} + \tfrac{1}{2}u_i''(x_i^{FB})\,\delta_i^2.$$

Therefore

$$W(t,n) \;=\; W^* + \tfrac{1}{2}\,\mathbb{E}\big[u_i''(x_i^{FB})\,\delta_i^2\big] - \kappa n\,\mathbb{E}[x_i^*]. \tag{14}$$

Lemma 1 (Wedge representation)**.** *Under the quadratic approximation,* $\delta_i = D_i'\,\Delta_i(t,n)$*, where* Δ_i *is the policy wedge in* (10)*.*

Proof. By (8) and (9),

$$u_i'(x_i^*) - u_i'(x_i^{FB}) = \big(p + \kappa n - (1-\theta n)b_i\big) - (c + \mathbb{E}[\xi]) = t + \kappa n - (1-\theta n)b_i - \mathbb{E}[\xi] = \Delta_i.$$

Under the quadratic approximation, $u_i'(x) = u_i'(x_i^{FB}) + u_i''(x_i^{FB})(x - x_i^{FB})$, so $\Delta_i = u_i''(x_i^{FB})\,\delta_i$. Using $u_i''(x_i^{FB}) = 1/D_i'$ gives $\delta_i = D_i'\,\Delta_i$. □

Substituting Lemma 1 into (14) and using $u_i''(x_i^{FB})\,(D_i')^2 = D_i'$,

$$\tfrac{1}{2}\,\mathbb{E}\big[u_i''(x_i^{FB})\,(D_i'\,\Delta_i)^2\big] \;=\; \tfrac{1}{2}\,\mathbb{E}\big[D_i'\,\Delta_i^2\big] \;=\; \tfrac{1}{2}\,\mathbb{E}[D_i']\,\mathbb{E}_W\big[\Delta_i^2\big]\,,$$

where the last equality uses the definition of $\mathbb{E}_W[\cdot]$ and the sign convention $D_i' < 0$. We have shown:

Proposition 1 (Welfare function)**.** *Under* (8)–(12)*,*

$$W(t,n) \;=\; W^* + \tfrac{1}{2}\,\mathbb{E}[D_i']\,\mathbb{E}_W\Big[\big(t + \kappa n - (1-\theta n)b_i - \mathbb{E}[\xi]\big)^2\Big] - \kappa n\,\mathbb{E}[x_i^*]. \tag{15}$$

Equivalently, the welfare loss relative to the first-best is

$$L(t,n) \;\equiv\; W^* - W(t,n) \;=\; -\tfrac{1}{2}\,\mathbb{E}[D_i']\,\mathbb{E}_W\big[\Delta_i(t,n)^2\big] + \kappa n\,\mathbb{E}[x_i^*] \;\geq\; 0.$$

Equation (15) is the Harberger-style quadratic loss representation: welfare falls below the first-best in proportion to the squared wedge between the policy-augmented price and the corrective target $(1-\theta n)b_i + \mathbb{E}[\xi]$, weighted by demand responsiveness. The second term captures the psychic cost of the nudge, which is destroyed rather than redistributed and so enters welfare with a unit coefficient.

E.3 Optimal Tax

Specialize Proposition 1 to the tax-only regime ($n = 0$):

$$W(t,0) \;=\; W^* + \tfrac{1}{2}\,\mathbb{E}[D_i']\,\mathbb{E}_W\left[(t - b_i - \mathbb{E}[\xi])^2\right].$$

The first-order condition with respect to t is

$$\frac{\partial W(t,0)}{\partial t} \;=\; \mathbb{E}[D_i']\,\mathbb{E}_W[t - b_i - \mathbb{E}[\xi]] \;=\; 0,$$

which yields the optimal Pigouvian tax

$$t^* \;=\; \mathbb{E}_W[b] + \mathbb{E}[\xi]. \tag{16}$$

The optimal corrective tax is the demand-slope-weighted average bias plus the expected externality. At $t = t^*$ the residual wedge is $b_i - \mathbb{E}_W[b]$: a uniform tax prices the mean of the distortion but cannot eliminate its dispersion. The residual welfare loss is therefore $-\frac{1}{2}\,\mathbb{E}[D_i']\,\mathrm{Var}_W(b)$, and the welfare gain at the optimum relative to the no-policy benchmark $(t = 0, n = 0)$ is

$$\Delta W_{\text{tax}} \;\equiv\; W(t^*,0) - W(0,0) \;=\; -\tfrac{1}{2}\,\mathbb{E}[D_i']\left(\mathbb{E}_W[b] + \mathbb{E}[\xi]\right)^2. \tag{17}$$

Equation (17) is the Harberger gain from setting a single corrective tax equal to the average distortion: it scales quadratically with the sum of the average internality and the expected externality.

F Welfare Evaluation Under Current U.S. Policy

The welfare results in the main text (Figure 6) evaluate every policy as a departure from a laissez-faire benchmark with no pre-existing taxes or subsidies. That benchmark isolates the conceptual value of each intervention in general, but it is not the right object for advising *current* policy in a market where some corrective policy is already in place. This appendix re-evaluates the same framework taking current U.S. policy as the status quo, as in Hendren and Sprung-Keyser (2020). Figure 14 repeats the two panels of Figure 6 with one change: each market's pre-existing per-unit tax or subsidy t_0 is set to its current U.S. value rather than to zero. Setting $t_0 = 0$ reproduces the estimates in the main text, so any difference is attributable to existing policy.[18]

Table 2 lists the policies we take as given. The cigarette tax combines the federal excise of \$1.01 per pack with the average state excise of about \$2.00. The alcohol tax combines the federal excise on beer, wine, and spirits with the corresponding average state excise per proof gallon, weights the three categories by their shares of U.S. ethanol consumption, and expresses the total per 10 g of pure ethanol (the WHO standard-drink unit), which puts the current tax at \$0.07 per 10 g. This is about \$0.10 per average US drink. For flu vaccines we model the status quo as a full subsidy equal to the dose price, because the Affordable Care Act, Medicare Part B, and the Vaccines for Children program make recommended immunizations free to most patients. Electricity carries no federal excise or carbon tax, so we use only the state and local taxes on residential elec-

[18]Unlike pre-existing taxes and subsidies, we do not need to account explicitly for pre-existing nudges. The baseline bias b in our framework is the bias prevailing in the current market, so it already embeds whatever changes in bias existing nudges (or marketing campaigns by firms) have accomplished. A nudge drawn from our sample then shifts this prevailing bias by some $(1 - \theta)b$, which is what we measure in the meta analysis.

tricity, about 3% of the bill. Residential water is sales-tax-exempt in most states.

Table 2: Current U.S. Policy and the Calibrated Optimum

Market	Current policy t_0	Optimum t^*
Energy	\$0.004/kWh	\$0.20/kWh
Cigarettes	\$3.01/pack	\$3.00/pack
Flu vaccines	−\$41/shot	−\$196/shot
Water	\$0/CCF	\$3.84/CCF
Alcohol	\$0.07/10 g	\$0.96/10 g

Notes: t_0 is the current per-unit tax in place in the US and $t^* = \mathbb{E}_W[b] + \mathbb{E}[\xi]$ is the calibrated optimum from the main text. Positive values are taxes; negative values are subsidies. Alcohol rates are expressed per 10 g of pure ethanol (the WHO standard-drink unit; a US standard drink is 14 g, so the figures are about \$0.10 and \$1.34 per US standard drink).

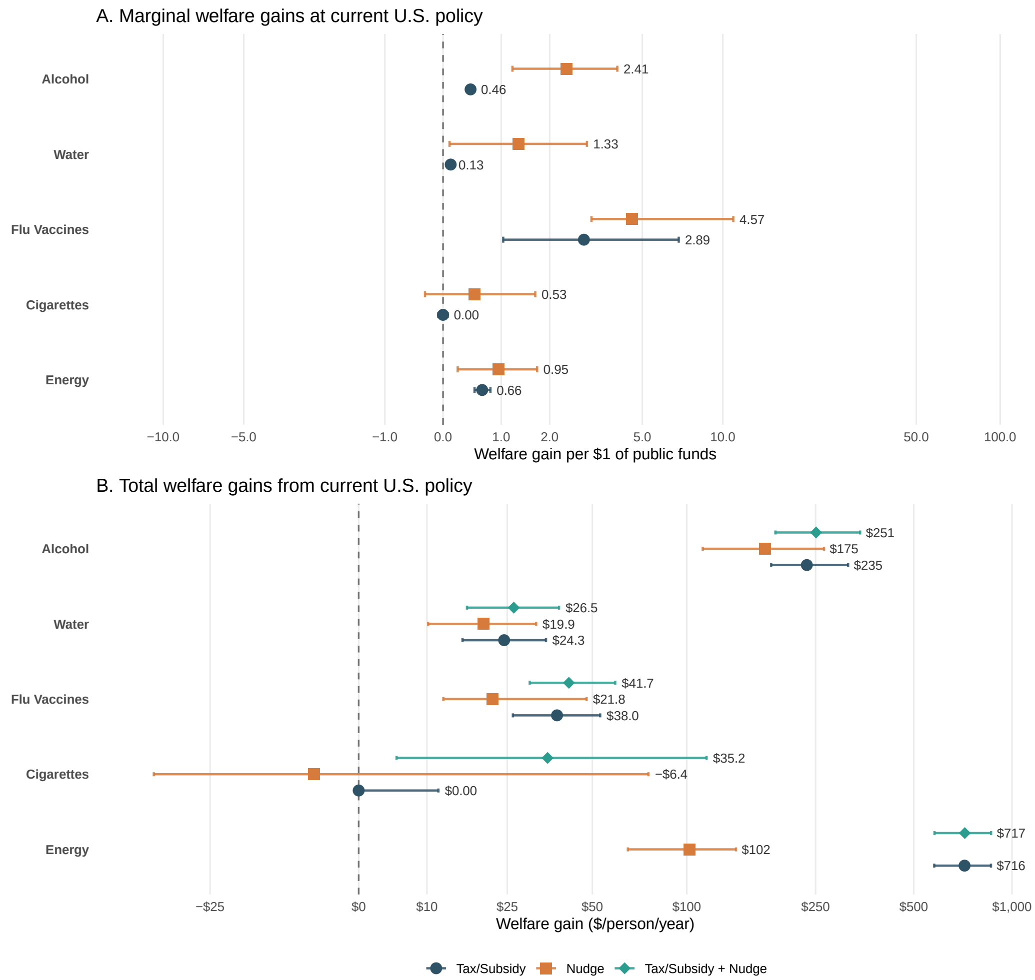
Taking current policy as given changes the picture meaningfully in only one market. The current cigarette tax of \$3.01 per pack is essentially equal to the calibrated optimum of \$3.00, so the entire corrective gain has already been realized. Of the \$69.2 per smoker per year that the optimal tax delivers relative to laissez-faire, \$69.2 is realized and essentially none remains. Because the existing tax already prices the externality and most of the average internality, an additional cigarette nudge now reduces welfare rather than raising it. Its marginal value falls from \$18.95 to \$0.53 per public dollar and its total effect from \$104 to $-\$6.4$ per smoker, as the nudge pushes consumption below the social optimum and erodes inframarginal tax revenue. For flu vaccines the gain is only partly realized. Free vaccination captures \$22.8 of the \$60.8 optimal-subsidy gain, but because the optimal subsidy of about \$196 per shot (a \$153 externality plus a \$43 internality) far exceeds the \$41 dose price, \$38.0 remains. The vaccine nudge thus stays valuable though less so than from a zero baseline, since each vaccination it induces also draws the existing subsidy.

In the remaining three markets the United States prices the relevant externality so weakly that current policy and laissez-faire deliver nearly the same welfare accounting.

The effective state and local electricity tax averages just \$0.004 per kWh—a national average heavily diluted by the widespread statutory exemption of residential electricity from state sales taxes. This baseline rate is two orders of magnitude below the optimal corrective tax of \$0.20 per kWh and captures only about 4% of the available gain. Residential water carries no corrective tax, so results are identical to the main text; and alcohol excise taxes of \$0.07 per 10 g (about \$0.10 per US standard drink) capture about 14% of the optimal gain. In all three markets, the ranking between nudges and taxes is unchanged from Figure 6.

The laissez-faire results in the main text are the right object for understanding the conceptual value of each instrument, while the status-quo results here are the relevant object for advising policy today. The two coincide except where existing corrective policy is already substantial, which among the markets we study means cigarettes and, to a lesser extent, influenza vaccines.

Figure 14: Marginal and Total Welfare Under Current U.S. Policy



Notes: This figure repeats Figure 6 taking current U.S. policy as the status quo rather than laissez-faire; the pre-existing tax and subsidy values t_0 are in Table 2. Panel A reports the marginal welfare gain per dollar of public funds ($1 - \mathrm{MVPF}_{\mathrm{tax}}$ for revenue raisers, $\mathrm{MVPF} - 1$ for spending policies). Panel B reports the total welfare gain from moving from current policy to the calibrated optimum, for the tax or subsidy alone, the nudge alone, and the two combined. All estimates use the baseline de-biasing calibration $\theta = 1$ and the same prices, quantities, elasticities, treatment effects, and externalities as the main text. Error bars are 95% confidence intervals from 2,000 bootstrap replications.