The Problem of Intra-Personal Cost

Brian Galle[*]

Abstract

"Externalities," or harms to others, provide a standard justification for government intervention in the private market. There is less agreement over whether government is justified in correcting "internalities," or harms to self the self is largely powerless to avoid. While some of the internality dispute is philosophical, some is practical. Critics suggest government lacks information to regulate internalities, and that any intervention would inefficiently distort a private market for self-help. In this Article, I show that these critiques of regulation overlook well-established tools of externality regulation, as well as a burgeoning literature on the measurement of internalities.

Having answered the "should" question, I move on to "how?". I examine the established tools of externality regulation, and consider to what extent the standard advice of the externality literature extends to internality regulation. I find some surprising results, such as that "carrots" may at times be an attractive alternative to "sticks," and that even large taxes on internalities can produce a so-called "double dividend." I also compare the traditional regulatory options to "nudges" and other forms of cognitively-informed government interventions. I show a set of cases in which nudges may be preferable to either taxes or command and control regulation.

Thus, my analysis also helps to resolve a second, related, debate over the propriety of nudges. The nudge debate has almost exclusively revolved around whether nudges avoid philosophical objections to paternalistic government regulation. I offer instead a new reason to employ nudges in some cases: they are more efficient.

Contents

Introduction

We humans are not always the best guardians of our own welfare. In the original movie version of *Total Recall*—let us not mention the forgettable remake—Arnold Schwarzenegger's character is on the hunt for the evil villain who has ruined Mars, only to discover that it was him all along.[1] The possibility that government might help us avoid these kinds of regrettable decisions is not new, dating back at least to Aristotle and possibly Homer, depending on how metaphorically one wants to read the Odyssey.[2]

But as vast and culturally pervasive as the "paternalism" literature has become, it has tended toward the philosophical, lingering on the propriety of government intervention to correct self-harms.[3] A decade, for instance, after Sunstein & Thaler first introduced the idea that government might "nudge" us toward better decisions,[4] the nudge debate remains caught up in cycles of argument over whether nudging is consistent with libertarian values.[5]

This is a frustrating state of affairs for those who are relatively comfortable with government intervention in the marketplace.[6] Human failings and new developments in how

---

[1] Whoops -- spoiler alert.

[2] *See* JON ELSTER, ULYSSES UNBOUND: STUDIES IN RATIONALITY, PRECOMMITMENT, AND CONSTRAINTS 3, 8 (2000).

[3] *E.g.,* JULIAN LE GRAND & BILL NEW, GOVERNMENT PATERNALISM: NANNY STATE OR HELPFUL FRIEND? 105–82 (2015).

[4] Cass R. Sunstein & Richard Thaler, *Libertarian Paternalism is Not an Oxymoron*, 70 U. CHI. L. REV. 1159, 1190–95 (2003); *see generally* RICHARD THALER & CASS SUNSTEIN, NUDGE (rev. & expanded ed. 2009).

[5] Cass R. Sunstein, *Behavioral Economics and Paternalism*, 122 YALE L.J. 1826, 1867–97 (2013); *Symposium*, *The Ethics of Nudging: Evaluating Libertarian Paternalism*, 14 GEORGETOWN J.L. & PUB. POL'Y (forthcoming 2016); Christian Coons & Michael Weber, *Introduction*, *in* PATERNALISM: THEORY AND PRACTICE 1, 15–23 (Christian Coons & Michael Weber eds., 2013).

[6] On Amir and Orly Lobel, *Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy*, 108 COLUM. L. REV. 2098, 2127–32 (2008); Ryan Bubb & Richard H. Pildes, *How Behavioral Economics Trims Its Sails and Why*, 127 HARV. L. REV. 1593, 1596–98 (2014); Lauren E. Willis, *When Nudges Fail: Slippery Defaults*, 80 U. CHI. L. REV. 1155, 1227–29 (2013). Bubb & Pildes argue that policy makers should "analyze [nudges] much as we would analyze explicit mandates," and make a "full comparison of the advantages and disadvantages of different regulatory instruments," *supra* at 1601, but they do not engage in that analysis themselves. This Article does.

The notable major exception to the absence of substantive analysis of internality regulation is a short recent policy brief focused on energy use, by Sunstein and Hunt Allcott, the NYU economist. Hunt Allcott & Cass R. Sunstein, Regulating Internalities, unpublished manuscript Feb. 2015, *available at* https://dash.harvard.edu/handle/1/16150609. Allcott and Sunstein briefly consider some of the issues I address, including the choice between different approaches to regulation. *Id.* at 7–9. But their treatment is cursory, omits most of the analysis offered here, and as a result goes astray at one or two points. *See infra* notes 219, 259.

Another partial analysis is Jacob Goldin & Nicholas Lawson, Nudges, Mandates, and Taxes: Policy Design with Active and Passive Decision-Makers, unpublished manuscript, Aug. 30, 2015, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2585038. Goldin & Lawson show persuasively that the combination of taxes and nudges can be superior to a flat ban on harmful choices along some dimensions, *id.* at 4–5, but they presume that nudges will always be used for "passive" actors, taxes for active. *Id.* at 13. My central question is which of these instruments is preferable. It also is unclear whether their framework extends to other irrationality settings. Finally, they do not consider income, revenue, or distributive effects, among other considerations examined here. *See infra* Part IV.

they can be addressed raise difficult and important questions, none so far addressed in the existing literature. Many governments already are deeply committed to helping consumers overcome what the governments perceive to be poor choices.[7] Canada, Australia, and many other countries around the world regulate tobacco with a complex regime in which manufacturers cannot display brand information, and instead must print disturbing images illustrating the health consequences of smoking.[8] The United States has proposed a similar policy, which currently is tied up in litigation.[9] Are scary pictures the best way to regulate smoking, or would something else, like a higher tobacco tax, be the best choice? Critics' qualms don't relieve courts and other actors in these regimes from having to confront the question of regulatory design.

Of course, the two questions are interrelated. A standard modern critique of government intervention, grounded in economic theory, is government's supposed inability to know better than the individual what will satisfy that person's preferences.[10] Even if government is right that there is a problem, compliance and enforcement carry costs that might make regulation wasteful overall.[11] If it could be shown that government interventions on average do better than private choices, though, this critique would weaken. And so perfecting regulatory design, the "how" question, also speaks to the "should" question.

Law and economics has already grappled with a similar set of problems in a similar context. A basic economic rationale for government regulation is the presence of externalities—harms or benefits that one of us creates for the others, in settings where the producer of the harms or benefits has limited incentives to care about the well-being of those affected by the spillovers.[12] Here, too, government faces the problem of incomplete information: how much does it cost to remedy an externality problem, and what is the value to society of the remedy?[13]

Over the past forty years or more, two related fields—law and economics, and environmental economics—have developed an elaborate set of answers to the challenges of limited information and transaction costs. Contemporary debates focus on a handful of key

---

[7] Coons & Weber, *supra* note 5, at 1.

[8] AUSTRALIAN GOVERNMENT DEPARTMENT OF HEALTH AND AGING, EVALUATION OF THE EFFECTIVENESS OF THE GRAPHIC HEALTH WARNINGS ON TOBACCO PRODUCT PACKAGING 12–15, http://webarchive.nla.gov.au/gov/20140801094920/http://www.health.gov.au/internet/main/publishing.nsf/Content/phd-tobacco-eval-graphic-health-warnings-full-report; David Hammond, *Health Warning Messages on Tobacco Products: A Review*, 20 TOBACCO CONTROL 327, 327 (2011).

[9] *R.J. Reynolds Tobacco Co. v. Food & Drug Admin*., 696 F.3d 1205, 1208 (D.C. Cir. 2012). The underlying First Amendment basis for the D.C. Circuit's initial rejection of the vivid images rules was later overruled by the Court sitting en banc in a different case, *Am. Meat Inst. v. U.S. Dep't of Agric*., 760 F.3d 18, 22–23 (D.C. Cir. 2014) (en banc), leaving the future of the rules unclear. Similar rules for smokeless tobacco are still in effect, *see* http://www.fda.gov/TobaccoProducts/Labeling/Labeling/ucm2023662.htm.

[10] Claire Hill, *Anti-Anti-Anti-Paternalism*, 2 N.Y.U. J. L. & LIB. 444, 445–48 (2007).

[11] Jeff Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 N'WESTERN UNIV. L. REV. 1165, 1219–25 (2003).

[12] JONATHAN GRUBER, PUBLIC FINANCE AND PUBLIC POLICY 4 (3d ed. 2011).

[13] *Id*. at 137–39.

design questions about the structure of regulation, and to a surprising degree have reached something like consensus on many points.[14]

My focus here will therefore be on to what extent the lessons of the externality-regulation literature apply to government efforts to protect us from ourselves.[15] Following the terminology of some leading economic commentators, I will call these failures of self-regard "internalities."[16] I first show how many lessons of the externality literature help to resolve practical complaints about whether we should even be engaged in paternalistic regulation, and document how more recent scholarly innovations go even further to resolve critics' concerns. I then move on to more concrete design questions.

To preview briefly my results, I find that many settled lessons of the externality literature are likely to be different, often profoundly different, in the internality context. "Command and control" regulation or its contemporary cousin, the "nudge," could dominate corrective taxation; rewards might be better than punishments, and legal rules can be important tools of redistribution. These points probably need some unpacking for those who aren't already deeply immersed in externality regulation.

Let me first try to be clear at the outset what I mean by an internality. What I have in mind is an outcome that the individual, if they deliberated about their choice in a coolly reflective, objective moment, would reject.[17] I wish I had gone to the gym last week, and that I hadn't eaten that second slice of pecan pie, and that I had saved more for retirement.

I want to distinguish these kinds of regretted outcomes from simple ignorance. Sometimes, we go wrong because we don't have all the information to make the right choice. Often, though, it's rational for us not to gather all the data ourselves, since information acquisition is costly.[18] In these cases it is relatively straightforward that government should just provide the information, or subsidize its production by others,[19] although to be sure the design of

---

[14] For an overview, see Brian Galle, *Tax, Command…or Nudge? Evaluating the New Regulation*, 92 TEX. L. REV. 837, 848–53 (2014).

[15] My focus on choice of instruments distinguishes this Article from the small handful of earlier efforts at analyzing paternalistic regulation through economic tools. In Eyal Zamir, *The Efficiency of Paternalism*, 84 VA. L. REV. 229 (1998), Prof. Zamir offers a model for deciding when regulation of individual failings will on net increase welfare, with the main factors being a balance of individual benefit against the "frustration" and administrative costs of regulating. *Id.* at 263–65. Zamir does not attempt to distinguish between different regulatory options.

[16] *E.g.*, Jonathan Gruber, *Tobacco at the Crossroads: The Past and Future of Smoking Regulation in the United States*, 15 J. ECON. PERSPECTIVES 193, 206 (2001). The term is generally attributed to Herrnstein et al., *Utility Maximization and Melioration: Internalities in Individual Choice*, 6 J. BEHAV. DECISION MAKING 149, 149 (1993).

[17] Zamir, *supra* note 15, at 237; see Allcott & Sunstein, *supra* note 6, at 12 ("[P]aternalistic regulation can be limited to situations in which individuals' choices are demonstrably inconsistent….").

[18] John Conlisk, *Why Bounded Rationality?*, 34 J. ECON. LITERATURE 669, 671 (1996); Roy Radner, *Bounded Rationality, Indeterminacy, and the Theory of the Firm*, 106 ECON. J. 1360, 1363 (1996).

[19] Alan Schwartz, *Regulating for Rationality*, 67 STAN. L. REV. 1373, 1375–76 (2015).

the best information-sharing regime is not always obvious.[20] Our challenge here is different: what should we do about people who might have the necessary information available, but still act or fail to act in ways that are wrong for them?

As we'll see, human decision making can go wrong in a number of different ways. I'll argue that the best regulatory design for a given problem often will be different, depending on what kind of error individuals are making. Thus, one of the analytic contributions of the paper will be to group and categorize these errors in ways that, I argue, are analytically useful.

With that definitional work out of the way for the moment, let's turn back to the externality-regulation literature. A central issue for any would-be regulator is to choose what tools or "instruments" the government will employ.[21] Should government regulate using "prices" or more traditional "command and control" regulation? If it's a price, should the price be a penalty (stick) or reward (carrot)? A carbon tax, a subsidy for going green, or a hard cap on the tons of carbon emitted? In recent work I suggest that behaviorally-informed policies with "surprising" impact, including what Sunstein & Thaler term "nudges," also can be fit into this framework.[22]

While there is some nuance in the literature's answers, as with any important and complex question, the general consensus is that sticks are the best choice.[23] The government can design a price to capture most of the features of a hard cap, and in addition sticks bring in money and elicit more information from the public.[24] Carrots can reveal private information, but also cost money, and even worse they may induce moral hazard: bad actors will commit bad deeds simply in order to be paid to stop.

I'll argue here that often none of these traditional advantages of sticks apply to internality regulation. Moral hazard is a small concern for internalities, because it is difficult to credibly threaten to injure oneself for gain. Prices induce rational actors to reveal their private costs of compliance with the government's preference, but of course the problem for internality sufferers

---

[20] For example, recent commentators critique most efforts to cure information market failures, OMRI BEN-SHAHAR & CARL SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE 59–118 (2014), and propose extensive (if pessimistic) design solutions for some of these problems, *id*. at 121–37.
  I don't mean to claim that there is always a clear-cut difference between simple ignorance and more complex cognitive problems. *See* Zamir, *supra* note 15, at 254. I only intend to rule out those cases of information failure that indeed are straightforwardly rational.
[21] Jonathan Baert Wiener, *Global Environmental Regulation: Instrument Choice in Legal Context*, 108 YALE L.J. 677, 755–60 (1999); Brian Galle, *The Tragedy of the Carrots*, 64 STANFORD L. REV. 797, 813–40 (2012).
[22] Galle, *supra* note 14, at 854–59. The main difference between that earlier paper and this is the earlier work focuses almost entirely on externalities, or in a few instances externalities mixed with internalities.
  The category of "surprising" regulation is broader than nudges. According to Sunstein & Thaler, a nudge is an instrument that, for a fully rational actor, would have no or minimal welfare effects. A surprising regulation, in my systematization, is any whose effects are larger than rational-choice theory would predict, and so can include instruments whose impact is non-negligible. *See* Ian Ayres, *Regulating Opt-Out: An Economic Theory of Altering Rules*, 121 YALE L.J. 2032, 2087 (2012) (noting that "sticky defaults" may have "moderate" cost).
[23] *See infra* Part I.A.
[24] *Id*.

is that their responsiveness does not necessarily reflect the real long-term costs and benefits they face. Government must use other methods, such as experiments, menus, and self-targeting regulation, to reach the right actors when it regulates internalities, and these methods work roughly the same whether the instrument is denominated in dollars or not.

Taxes and other sticks still bring in more money than other options, of course, but in some instances the revenues are not worth the extra costs they bring. Building on my earlier work in the externality context, I show that "nudges" can potentially be a better choice on net, despite the absence of revenue. On this front I depart from earlier work by economists arguing that "small" internality-correcting taxes on soda or other tempting foods can be highly efficient.[25] I argue that this claim does not always hold for taxes large enough to affect consumers' labor-supply decisions.

Overall, it will turn out that the best choice of instrument depends on what kind of mistake individuals are making. For those of us who struggle with willpower or impatience, taxes may dominate. For those of us who go wrong through failures of attention—for instance, by neglecting our retirement savings—nudges look, at least on current available evidence, to be a more promising alternative.

In sum, while the lessons of the externality literature may not apply fully in the internality context, the analytical tools of that literature remain powerful. Even at this early stage of academic study of internalities, we can form some good hypotheses about what an efficient internality-regulation regime might look like. Whether that is enough to satisfy critics of internality regulation I cannot say. But many governments are already embarked on fairly extensive internality regulation. My analysis here offers a first, tentative, glance at how those regulations should take shape in the future.

I. Background and Prior Literature

This Part provides context for readers who may be unfamiliar with aspects of my argument. Part I.A. offers a general overview of economic approaches to externality regulation.[26] Part I.B. then briefly summarizes the concept of "internalities." Finally, Part I.C. offers some definitions and caveats going forward; even readers familiar with the concepts in Parts I.A. and I.B. may want to briefly visit I.C.

A. Regulating Externalities

Modern economic theories of government regulation begin with the premise that markets sometimes fail.[27] Externalities are a classic example.[28] An externality, simply put, is a harm

---

[25] Ted O'Donoghue & Matthew Rabin, *Optimal Sin Taxes*, 90 J. PUB. ECON. 1825, 1827 (2006).

[26] Part I.A. repeats, in essentially identical form, my earlier summary of this topic. Galle, *supra* note 14, at 843–46.

[27] GRUBER, *supra* note 12, at 3.

[28] *Id.* at 4.

("negative externality") or benefit ("positive externality") that affects someone other than the actor making an economic decision.[29]

In general, the goal of regulation is neither to eliminate negative nor to produce boundless quantities of positive externalities, but rather to achieve what might be called the optimal level of externality.[30] Eliminating even the worst pollutants is costly. Should government bankrupt coal producers, or is there a way to balance clean air against the costs of achieving it? On the positive externality side, everyone might agree that charity is beneficial. But should government spend millions to clothe or educate one more child?

Economists typically answer these kind of balancing questions using marginal analysis.[31] Under this approach, the policy maker asks herself, "on the margin---that is, for the very next unit of good or bad produced---what is the harm or benefit of that one unit for *everyone in society*?" We might therefore call this the "marginal *social* damage," (MSD) in the case of a negative externality, and "marginal social benefit" (MSB) for a positive one. She then compares this harm or benefit against the marginal costs to the producer. If the producer's private marginal cost is greater than the marginal social damage, it doesn't pay, on net, to prevent the damage: counting the producer's losses, society would lose by forcing the producer to avoid the externality.[32]

To see this graphically, consider Figure One.

---

[29] *Id*. at 122--23.

[30] *Id*. at 137--39; Gloria E. Helfand et al., *The Theory of Pollution Control*, in 1 HANDBOOK OF ENVTL. ECON. 249, 253 (Karl-Goran Maher & Jeffrey R. Vincent eds., 2003).

[31] GRUBER, *supra* note 27, at 126.

[32] Note, importantly, that for simplicity we are assuming here that we should count the costs and benefits for the producer and everyone else equally. That's a controversial proposition, but I'll leave it aside here for ease of exposition.

Figure One: Optimal Externality Production

Cost

Producer's
Marginal
Costs

A

τ

Social Marginal
Damage / Benefit

More Charity

More Pollution

Quantity
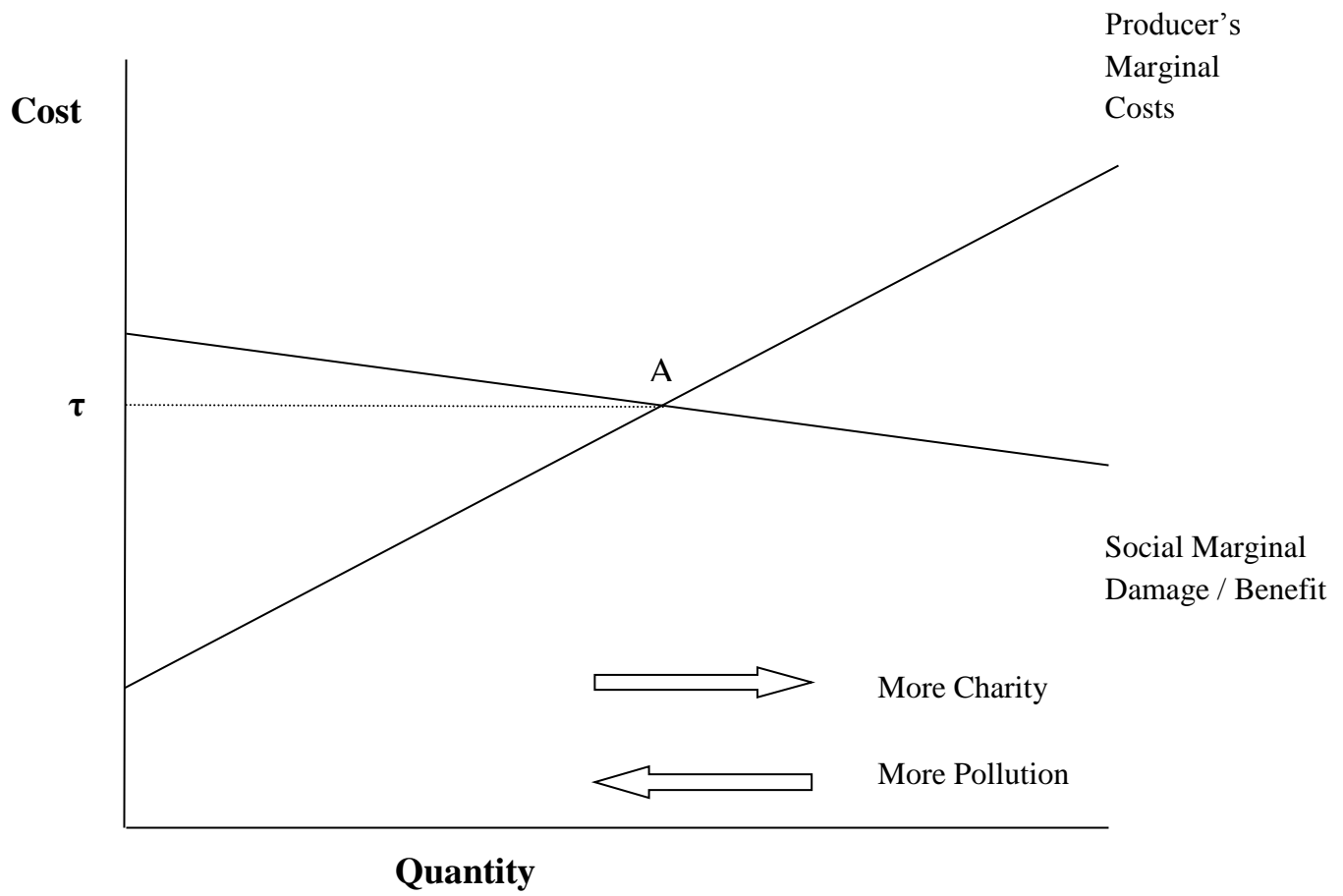
In figure one, the upward-sloping line represents the marginal cost curve for the externality producer: as we trace the line rightwards, each additional unit of pollution reduction (say, one ton less of carbon) or charitable output (say, another bed in a homeless shelter) is costlier to achieve.[33] The downward-sloping line is the marginal social benefit curve: each unit is slightly less beneficial than the last.[34] At point A, the two lines intersect. This is the optimal point.[35] Anywhere to the right of A, the costs of charity or pollution reduction outweigh the benefits. To the left, we've left cost-effective improvements on the table.

We could imagine a few ways of achieving production at this level A. If government knew the shapes of the two curves, it could calculate the quantity of output at A and simply mandate that producers achieve it, with jail for those who refuse.

Another approach is to set a price for producers. In the case of pollution, government could impose a fee or tax on each unit of carbon, in an amount equal to the producer's marginal cost at point A; this price is labeled "tau" in Figure One. For producers whose costs of eliminating the next unit of carbon are below tau, they will eliminate it, saving themselves tau minus their cost. For producers whose costs are above tau, they will simply emit the carbon and pay the tax. Thus, just as with the mandate, rational producers should produce exactly the amount of carbon at point A. Or, similarly, government could pay producers to eliminate carbon or produce charity. Once more, if the government offers a price tau, only producers who can fill a shelter bed for less than tau will take the offer.

Economists often call the first of these approaches "quantity regulation,"[36] and the second two "price instruments."[37] Lawyers may be more familiar with the similar divide between what Calabresi and Melamed termed "property rules" and "liability rules."[38]

Most commentators strongly favor price instruments over quantity regulation, except in settings where special administrative considerations make prices impractical.[39] As Kaplow &

---

[33] This reflects the likelihood that firms will undertake the cheapest efforts first, and then have to work harder and harder to achieve further milestones. For instance, at some point, adding more beds means building a new building.
[34] Again, diminishing marginal utility is a standard assumption here. We probably house the neediest persons first, and at some point we're offering shelter space to Bill Gates.
[35] I'm simplifying here for the sake of exposition. A more rigorous approach to setting the optimal quantity would also account for other factors that might affect the efficiency of the regulation. For example, if the regulation imposes costs, and the expectation of those costs changes behaviors other than the production of the externality—for example, distorts consumer choices among products—the ideal regulation might balance disruption of these expectations against pollution control. Helmuth Cremer et al., *Externalities and Optimal Taxation*, 70 J. PUB. ECON. 343, 346 (1998).
[36] GRUBER, *supra* note 12, at 137.
[37] STERNER, *supra* note **Error! Bookmark not defined.**, at 214--15.
[38] Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972); *see also* Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713, 715 (1986).
[39] GRUBER, *supra* note 12, at 140; Don Fullerton et al., *Environmental Taxes*, *in* DIMENSIONS OF TAX DESIGN: THE MIRRLEES REVIEW 231 (James Mirrlees ed. 2011); Maureen L. Cropper & Wallace E. Oates, *Environmental Economics: A Survey*, 30 J. ECON. LIT. 675, 686 (1992); Cameron Hepburn, *Regulation by Prices, Quantities, or*

Shavell show, prices can be used to duplicate most of the features of mandates.[40] Prices provide vital information to the government that regulation supposedly does not.[41] Further, prices are said to provide for revenues that the government can use for other projects.[42]

Glaeser's critique of "nudges" and similarly novel behavioral forms of regulation is typical in this regard. Glaeser argues that both taxes and nudges create economic distortions, but only taxes bring in money to help offset those losses.[43] Perhaps graphic images of the harms of cigarette smoking printed on the sides of packs would be repulsive enough that the smokers switch to cigarillos or pipe tobacco, which are nearly as harmful but which they enjoy less. Or perhaps some smokers cannot quit, but also suffer added pain as a result of the imagery.[44] Further, unlike a cigarette tax, the graphic images don't bring in any revenues that could be used to improve the lives of smokers or anybody else. To take another example, workers who are defaulted into a savings program, who are unwilling to pay the costs of the opt-out mechanism, and who genuinely would prefer not to save, are worse off than in the absence of the nudge.[45] A number of other economists have recently made a similar point about the preferability of a carbon tax over other regulatory alternatives: regulations change consumption patterns, creating deadweight loss, but bring in no offsetting dollars.[46]

B. Internalities

Harms done to others are a classic economic rationale for government regulation, but what about harms done to self? Most readers likely know that a large body of literature now suggests that individuals make decisions—or fail to make them—in ways that in the long run likely do not maximize their own subjective well-being. [47] Some commenters, seizing on the

Both: A Review of Instrument Choice, 22 OXFORD REV. ECON. POL'Y 226, 228--29 (2006). As an example of a "special consideration," price instruments may be riskier than quantity regulation when the marginal social damage curve is steep but its exact shape is uncertain, GRUBER, *supra* note 27, at 143--46, and the policy maker cannot sharply vary the tax rate to account for this risk.

[40] Louis Kaplow & Steven Shavell, *On the Superiority of Corrective Taxes to Quantity Regulation*, 4 AM. L. & ECON. REV. 1, 7–10 (2002).

[41] *Id.* at 4.

[42] *E.g.*, Helfand et al., *supra* note **Error! Bookmark not defined.**, at 287.

[43] Glaeser, *supra* note **Error! Bookmark not defined.**, at 135, 150; *see also* Korobkin, *supra* note **Error! Bookmark not defined.**, at 1669; Mitchell, *supra* note **Error! Bookmark not defined.**, at 1268, 1274; Rizzo & Whitman, *supra* note **Error! Bookmark not defined.**, at 960--61.

[44] Andrew Caplin, *Fear as a Policy Instrument*, *in* TIME AND DECISION: ECONOMIC AND PSYCHOLOGICAL PERSPECTIVES ON INTERTEMPORAL CHOICE 441, 452 (George Loewenstein et al. eds., 2003); *see* Fennell, *supra* note 49, at 1415 (making this point about imperfectly targeted nudges generally).

[45] Jeff Rachlinski, *Cognitive Errors, Individual Differences, and Paternalism*, 73 U. CHI. L. REV. 207, 224--25 (2006).

[46] Helfand et al., *supra* note **Error! Bookmark not defined.**, at 287; Ian Parry et al.*, When Can Carbon Abatement Policies Increase Welfare? The Fundamental Role of Distorted Factor Markets*, 37 J. ENVTL. ECON. & MGMT. 52, 52 (1999).

[47] For reviews, see B. Douglas Bernheim & Antonio Rangel, *Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision Makers*, *in* BEHAVIORAL ECONOMICS AND ITS APPLICATIONS 7, 10–65 (Peter Diamond & Hannu Vartiainen eds., 2008); Stefano DellaVigna, *Psychology and Economics: Evidence from the Field*, 47 J. ECON. LITERATURE 315 (2009).

externality analogy, have dubbed these kinds of mistakes "internalities": costs that the deciding self inflicts on its temporal successors.[48]

Because a good deal of my later discussion will turn on the details of how humans go wrong, it's worth highlighting some aspects of the empirical literature here. One key finding is that we are overwhelmingly creatures of the present, and only through exercises of our limited pool of willpower can we force ourselves to take sufficient account of the future.[49] Relatedly, we tend to focus our attention on facts that are readily available to us or on items in plain sight, reacting automatically and emotionally to those immediate stimuli.[50] The Nobelist Daniel Kahneman calls these two modes of reasoning, the unconscious and the deliberative, "system one" and "system two," respectively. Only with some effort do we turn our attention to the distant and the hidden, and engage our system two reasoning powers to reach better decisions.[51] We "anchor" on information we've already received, and interpret new data selectively to fit with what we already know or want to be true.[52] In all of these areas evidence suggests that individuals vary considerably in their susceptibility to the behavior.[53]

The consequences of these human tendencies can be seen all around us. Few human institutions, from families up through the federal government, make adequate plans for their financial future.[54] We procrastinate or give in to temptation, then build costly structures to overcome our tendencies, and then incur even more costs to unwind them.[55] People smoke too much, don't exercise enough, eat to excess. Many of us, even trained experts, make decisions based on only a fraction of the information available to us, choosing poor investments and neglecting "hidden" costs that in actuality are easily calculable.[56]

Importantly for my later analysis, evidence so far suggests that some of us are more self-aware of these failings than others. Consider the example of the mutual bank. Mutuals offer credit cards with relatively higher interest rates but promise "no hidden fees."[57] That combination of features seems most plausibly aimed at customers who know their own tendency

---

[48] *See supra* note 16.

[49] For overviews of the literature, see Lee Anne Fennell, *Willpower Taxes*, 99 GEO. L.J. 1371, 1375--94 (2011); Shane Frederick et al., *Time Discounting and Time Preference: A Critical Review*, *in* ADVANCES IN BEHAVIORAL ECONOMICS 162, 166--79 (Colin F. Camerer et al. eds., 2007).

[50] Daniel Kahneman, *Maps of Bounded Rationality: Psychology for Behavioral Economics*, 93 AM. ECON. REV. 1449, 1451–57 (2003).

[51] *Id*. at 1467--69.

[52] JONATHAN BARON, THINKING AND DECIDING 203--24, 263--70 (4th ed. 2008)

[53] Ted O'Donoghue & Matthew Rabin, *Self-Awareness and Self-Control, in* TIME AND DECISION: ECONOMIC AND PSYCHOLOGICAL PERSPECTIVES ON INTERTEMPORAL CHOICE 217, 219--20 (George Loewenstein et al. eds., 2003).

[54] Shlomo Benartzi & Richard Thaler, *Heuristics and Biases in Retirement Savings Behavior*, 21 J. ECON. PERSP. 81, 82–84 (2007).

[55] Frederick et al., *supra* note 49, at 172--79.

[56] John R. Graham & Campbell R. Harvey, *The Theory and Practice of Corporate Finance: Evidence from the Field*, 60 J. FIN. ECON. 187, 188–243 (2001); Aradhna Krishna et al., *A Meta-Analysis of the Impact of Price Presentation on Perceived Savings*, 78 J. RETAILING 101, 101--18 (2002).

[57] Ryan Bubb & Alex Kaufman, Consumer Biases and Mutual Ownership, NYU Law & Economics Working Paper 11-35, at 18--27 (Oct. 2011), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1945852.

to fall for the tricks played by other banks.[58] Mutuals command a small sliver of the credit market, however. Similarly, many households report that they let the government keep too much in tax withholding each year so that they will not face the temptation to spend that money too soon---and then, ironically, some of these same households later pay very high fees to get access to their money a few weeks early.[59] Though other interpretations are possible, a reasonable inference is that our understanding of our own frailty, even if present, is often imperfect.]

C. Philosophical Foundations & Other Clarifications

Although there now is extensive evidence that individuals make decisions that don't satisfy their own long-run preferences, there has been little scholarly analysis of how best to remedy that problem. Debate instead is stuck at a more fundamental question: should government be in the business of correcting internalities at all? Critics assert that government intervention is unwarranted "paternalism."[60] Behind the paternalism label are two deeper critiques: that governments lack the capacity to regulate in ways that will lead to better outcomes, and that humans should have the autonomy to make their own mistakes.[61]

The autonomy argument poses difficult philosophical problems that law & economics lacks the tools to resolve. For me, law that helps individuals achieve their true goals furthers autonomy, rather than undermining it. That, after all, is the structure of constitutions: they protect bodies politic from momentary whims and passions, and preserve the capacity for long-run self-determination.[62] But I recognize that some readers will have philosophical commitments that make it hard for them to accept this claim.

Economics has more to offer in the debate over whether government has the capacity to regulate internalities. I will try to resolve these questions in Part III. The answers, we'll see, often depend on just how individuals are failing themselves. Therefore, we will first need a brief taxonomy of internalities; Part II takes up that task.

Before that, it is worth offering a few short clarifications to the scope of my task. One is that, following the prevailing law & economics literature approach to cost-benefit analysis, my approach is essentially welfarist.[63] I am interested in which set of rules maximizes total social welfare, assuming diminishing marginal utility and some degree of popular preferences for

---

[58] *Id.* at 4.
[59] Michael S. Barr & Jane K. Dokko, *Third-Party Tax Administration: The Case of Low- and Moderate-Income Households*, 5 J. EMPIRICAL L. STUDIES 963 (2008).
[60] Richard A. Epstein, Exchange, *The Neoclassical Economics of Consumer Contracts*, 92 MINN. L. REV. 803, 806–07 (2008); Heidi M. Hurd, *Fudging Nudging: Why Libertarian Paternalism is the Contradiction It Claims It's Not*, 15 GEORGETOWN J.L. & PUB. POL'Y __ (forthcoming 2016), manuscript at 8.
[61] Epstein, *supra* note 60, at 806–07; Joshua D. Wright & Douglas Ginsburg, *Behavioral Law & Economics: Its Origins, Fatal Flaws, and Implications for Liberty*, 106 NORTHWESTERN L. REV. 1033, 1065–74 (2012); Coons & Weber, *supra* note 7, at 7–9.
[62] Frederick Schauer, *Judicial Supremacy and the Modest Constitution*, 92 CAL. L. REV. 1045, 1054–55 (2004); *see* NORMAN DANIELS, JUST HEALTH CARE 159 (1985) (offering this rationale as a justification for so-called paternalistic regulation, despite alleged autonomy concerns).
[63] Zamir, *supra* note 15, at 233–35.

distributive fairness.[64] Of course, there could be alternative consequentialist approaches to these same questions, such as the suggestion by Sen and Nussbaum that we maximize basic human capabilities,[65] or perhaps a Rawlsian-inspired approach that would maximize health over other outcomes.[66] We also could consider deontological approaches, such as one that prioritized autonomy or dignity.[67] I don't mean to suggest these approaches are invalid, but they aren't in common use in the externality literature. My goal here is to focus first on translating what we already know about externalities to internalities.

Second, it might be objected that some or even most internality problems actually present a mix of internalities and externalities.[68] What, then, does an internality analysis add? My answer is that we can think of the internality as offering a reason for more extensive regulation. We just saw that the optimal level of regulation depends on the marginal social damage of a product.[69] We should include both externalities and internalities in calculating the marginal social damage.[70] This can make a dramatic difference in the regulator's choices. For instance, Gruber and Koszegi estimate that the back of the envelope externality cost of a pack of cigarettes is less than $1, while the internality cost is more like $30.[71] If government currently imposes only a small per-pack cost on smoking, its decisions about whether and how to correct internalities may determine if it will now impose a massive tax hike.

Finally, I emphasize that my analysis is aimed at genuine failures of decision, not simply ignorance of the best choice. Because information often has many of the features of a public good—that is, my investment in acquiring information produces positive externalities for others I usually cannot charge them for—the economic case for government support for information creation is straightforward.[72] Likewise, when individuals can rely on informed others to act for

---

[64] Part III.A. discusses the challenge of measuring welfare when individuals' observable choices don't necessarily reflect their long-term preferences.

[65] MARTHA C. NUSSBAUM, FRONTIERS OF JUSTICE: DISABILITY, NATIONALITY, SPECIES MEMBERSHIP 69–81 (2006); AMARTYA SEN, INEQUALITY REEXAMINED 39–55 (1992).

[66] DANIELS, *supra* note 62, at 42–47; Lawrence O. Gostin, *Securing Health or Just Health Care? The Effect of the Health Care System on the Health of America*, 39 ST. LOUIS U. L.J. 7, 13 (1994).

[67] Wright & Ginsburg, *supra* note 61, at 1068–75; Hurd, *supra* note 60, at 2.

[68] *See* Wendy Mariner, *Paternalism, Public Health, and Behavioral Economics: A Problematic Combination*, 46 CONN. L. REV. 1817, 1833 (2014); *cf.* Katherine Pratt, *A Constructive Critique of Public Health Arguments for Antiobesity Soda Taxes and Food Taxes*, 87 TULANE L. REV. 73, 77–103 (2012) (examining both externality and internality rationales for obesity regulation).

[69] *See supra* note 31–36.

[70] Zamir, *supra* note 15, at 278. To be precise, the government should include the gap between the cost of the amount of internality the individual will consume on their own, and the cost of the optimal level of consumption, in the social marginal cost. In other words, if some consumers can avoid the internality on their own, but not completely avoid it, the price can be lower.

Admittedly, this approach does lead to potential difficulties if a given activity produces *positive* externalities and negative internalities. I reserve that case for future development.

[71] Jonathan Gruber & Botond Koszegi, *Is Addiction "Rational"? Theory and Evidence*, Q.J. ECON. 1261, 1291 (2001).

[72] Janusz Ordover & William Baumol, *Antitrust Policy and High-Technology Industries*, 4 OXFORD REV. ECON. POL'Y 13, 14 (1988). On the general theory of public goods, see RICHARD CORNES & TODD SANDLER, THE THEORY OF EXTERNALITIES, PUBLIC GOODS, AND CLUB GOODS 8 (2d ed. 1996).

them, we have "rational ignorance," supplying at least a basic argument in favor of regulation.[73] Having said that, human limits in absorbing and processing information can form an obstacle to good decisions even when the underlying information is freely available and individuals have incentives to employ it.[74]   These kinds of failures are within the scope of my analysis here.

II.   Mapping Internalities

To simplify exposition, we can think of individuals' attitudes towards internalities as the product of a two-by-two grid. To distinguish internality regulation from the (in my view) easier case of government information production, let us assume that full information about the costs and benefits of a given choice are freely available to all decision makers.

Let the first dimension of the grid, $\hat{\alpha}$, represent the marginal compliance cost perceived by the individual at the time of compliance. I will assume for present purposes that this subjective cost of compliance should properly be included in the social welfare function. Thus $\hat{\alpha}$ is the equivalent, in the internality context, of the private marginal cost faced by an externality producer for each unit of externality.
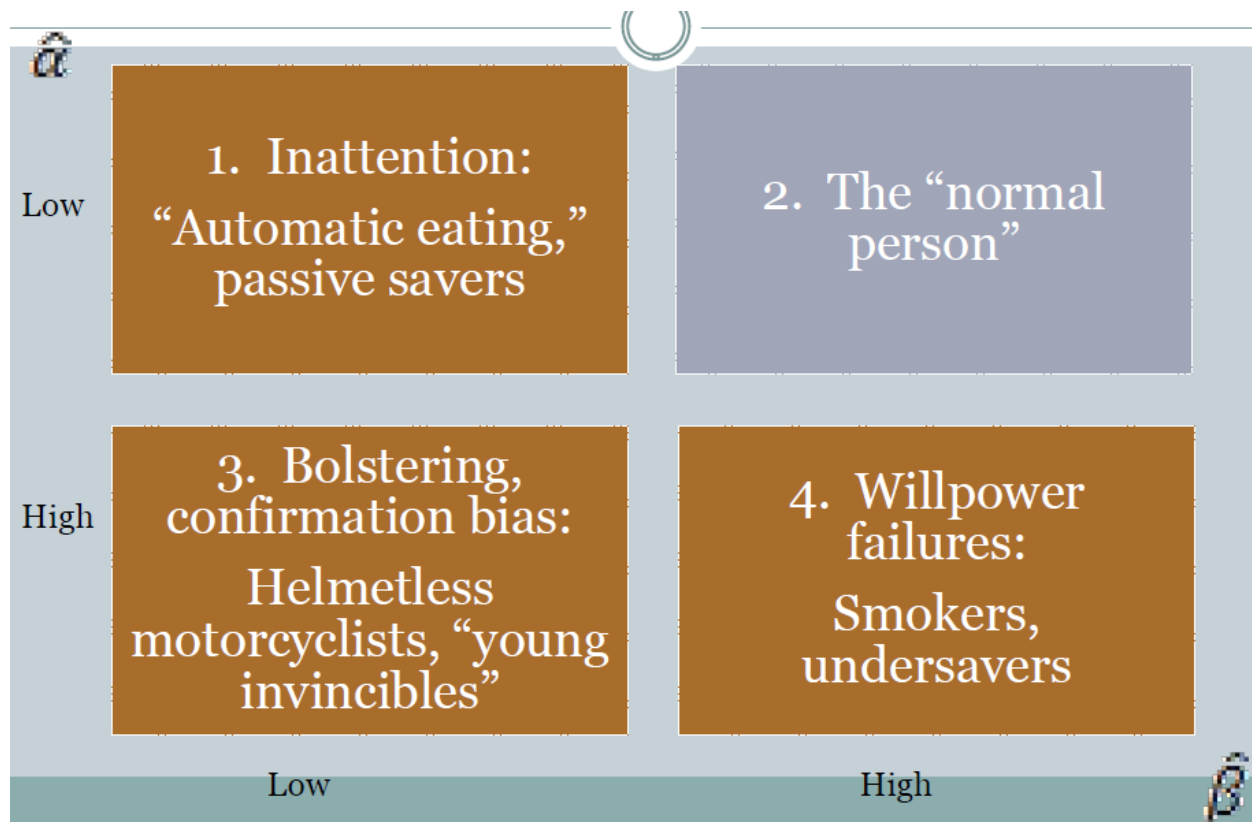
The second dimension of the grid can be represented by $\hat{\beta}$, or the subjective benefit the individual perceives at the time she must make the decision to comply or not. Since compliance or not determines whether the benefit occurs, $\hat{\beta}$ is also the value the individual assigns to a given government-preferred outcome at the time she makes the decision that contributes to that outcome. Also for simplicity, for now we will assume that the government's preference in fact would improve the individual's welfare relative to non-compliance. Under that assumption, $\hat{\beta}$ measures an individual's ability to perceive that her current decisions may not maximize her overall well-being. For example, she may recognize that she has a propensity to overvalue the future.

Figure Two illustrates the resulting possibilities, along with some illustrative examples.

---

[73] Howard Beales et al., *The Efficient Regulation of Consumer Information*, 24 J.L. & Econ. 491, 495–501 (1981); *see* Frank H. Easterbrook & Daniel R. Fischel, *Mandatory Disclosure and the Protection of Investors*, 70 Va. L. Rev. 669, 681–85 (1984) (explaining this point, but also arguing that private market can substitute for government encouragement in some cases).

[74] *See, e.g*., Julie S. Downs et al., *Strategies for Promoting Healthier Food Choices*, 99 AM. ECON. REV. (Papers & Proceedings) 159, 160—62 (2009) (summarizing two studies in which disclosing health information had no or perverse effects).

α

Low

1. Inattention: "Automatic eating," passive savers

2. The "normal person"

High

3. Bolstering, confirmation bias: Helmetless motorcyclists, "young invincibles"

4. Willpower failures: Smokers, undersavers

Low   High

β

Box 2, the upper right square of the grid, is the baseline case, the "normal" person.[75] She perceives the costs of complying with the government's policy as low and the benefits as high. In all likelihood, she needs no additional incentive to follow the government's course.

Individuals in the upper left box, Box 1, don't view the government's suggestion as burdensome, but also don't recognize that its goals are worthwhile. This box might capture the empirically well-documented phenomena of "inattentive" actors and "reference dependence," or the importance of framing and presentation on how we decide.[76] For example, Wansink and others find that portion size strongly influences many people's consumption of food and beverages; we eat what's in front of us, without really paying attention to how much we're eating.[77] Chetty et al. report evidence that 85% of Danish working households were unresponsive to tax incentives for savings, and also did not respond to changes in the default

---

[75] *But see* The Arcade Fire, Normal Person, on *Reflektor* (Arista 2013) ("I've never really ever met a normal person.").

[76] Kahneman, *supra* note 50, at 1451–57. For more in-depth reviews, see B. Douglas Bernheim & Antonio Rangel, *Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision Makers*, *in* BEHAVIORAL ECONOMICS AND ITS APPLICATIONS 7, 10–5 (Peter Diamond & Hannu Vartiainen eds., 2008); Stefano DellaVigna, *Psychology and Economics: Evidence from the Field*, 47 J. ECON. LITERATURE 315, 324–36, 347–56 (2009).

[77] BRIAN WANSINK, MINDLESS EATING 17–19, 47–52 (2006); Pierre Chandon, *How Package Design and Packaged-Based Marketing Claims Lead to Overating*, 35 APPLIED ECON. PERSPECTIVES & POL'Y 7, 13–18 (2013).

amount of savings the government chose for them.[78] In Kahneman's terminology, these are "decisions" that are made using system one alone.[79]

Box 4, the lower-right box, may capture willpower failures, another set of well-known behaviors.[80] Individuals know what's good for them, but in the moment they must make their decision, they find the bad choice too difficult to resist.[81] We should expect that high-$\hat{\alpha}$, high-$\hat{\beta}$ individuals will seek out "commitment devices," or tools to help them obtain the beneficial outcome.[82] Whether government intervention is justified to assist these households may depend on the extent to which commitment devices are unobtainable, create costly side-effects, or have unwanted distributive impacts.[83]

Individuals who suffer from various forms of "bounded rationality" may fall somewhere near the corners of Box 1 and Box 4.[84] Oftentimes we face problems we lack the cognitive capacity to absorb.[85] To economize on time and brainpower, we may take mental shortcuts that lead us to imperfect answers.[86] Accepting the government's choice might save us from having to make our own decision, but if so we don't have a good way to know if it's the best choice for us. Verifying that the government's suggestion is a good one would be somewhat costly, but also establish it as somewhat valuable. Also, in some cases it appears that the use of shortcuts is itself motivated by procrastination, making these cases a true hybrid.[87]

The last, lower-left, box is perhaps the most puzzling. Despite the presence of (by hypothesis) full information, individuals here largely ignore the benefits of the government's choice. Remember that we are assuming for now that the government's choice is correct; it isn't simply that individuals in this corner know their own tastes better. Nonetheless, since they perceive the cost of changing their own decision as high, they resist government proposals. This

---

[78] Raj Chetty et al., *Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark*, 129 Q.J. ECON. 1141 (2014).

[79] Kahneman, *supra* note 50, at 1451–57.

[80] For surveys, see Lee Anne Fennell, *Willpower Taxes*, 99 GEO. L.J. 1371,1375–94, and Shane Frederick et al., *Time Discounting and Time Preference: A Critical Review*, *in* ADVANCES IN BEHAVIORAL ECONOMICS 162, 166–79 (Colin F. Camerer et al. eds., 2004).

[81] *Id.*

[82] Gruber & Koszegi, *supra* note 71, at 1278; Ted O'Donoghue & Matthew Rabin, *Doing It Now or Later*, 89 AM. ECON. REV. 103, 105 (1999).

[83] Brian Galle & Manuel Utset, *Is Cap-and-Trade Fair to the Poor? Shortsighted Households and the Timing of Consumption Taxes*, 79 GEO. WASH. UNIV. L. REV. 33, 76–80 (2010).

[84] Herbert Simon, *Invariants of Human Behavior*, 41 ANN. REV. PSYCHOL. 1, 6 (1990).

[85] Conlisk, *supra* note 18, at 672. Bubb & Pildes, *supra* note 6, at 1613–14, summarize the evidence in the context of retirement savings.

[86] RICHARD THALER, QUASI-RATIONAL ECONOMICS 3–5 (1994); Kahneman, *supra* note 50, at 1458–59.

[87] Ted O'Donoghue & Matthew Rabin, *Procrastination in Preparing for Retirement*, *in* BEHAVIORAL DIMENSIONS OF RETIREMENT ECONOMICS 125, 125–26 (Henry J. Aaron ed., 1999); Kahneman, *supra* note 50, at 1468.

might describe the so-called "naïve hyperbolic discounters," those who are impatient but fail to recognize their own impatience.[88]

This set of behaviors could also reflect what is sometimes called "bolstering," or "cultural cognition."[89] That is, we tend to selectively filter our understanding of the world in ways that reinforce our preferred outcome.[90] Since the present self wants to feel the wind in its hair, it screens out or rejects as "biased" evidence that helmetless motorcyclists die in remarkable numbers.[91] Often the illusion of personal control is an important tool in the process of self-deception, allowing the actor to distinguish her own case from the statistical evidence: *other* people might die of lung cancer, but *I* can quit whenever I want.[92] A number of commentators point to these kinds of processes as also explaining why many households do not adequately insure against flood and other disasters.[93]

Again, the taxonomy is meant to be simplifying, and so it likely misses some relevant nuance. The boxes are not meant to represent points, but continua. Nonetheless, it is a starting place, and I will now argue that it at least in part reflects important differences between different potential sets of internalities.

III. Should We Regulate Internalities?

Before we proceed to design issues, I expect that many readers likely want some additional convincing that government should regulate internalities at all. Again, basic philosophical objections are beyond my scope here. The core of many putatively philosophical claims, however, are actually quite practical. For example, many self-identified libertarian thinkers object to paternalistic regulation because, they claim, the government cannot know individual preferences accurately enough to regulate effectively.[94] Thus in Part III.A. I will address the core of this informational argument against regulating internalities. A second common libertarian objection to any regulation is that the possibility of private ordering solutions make government action unnecessary. Part III.B. therefore considers whether government regulation of internalities would interfere with an efficient private market in internality correction.

---

[88] *See generally* Ted O'Donoghue & Matthew Rabin, *Choice and Procrastination*, 116 Q.J. ECON. 121 (2001) (developing a model of partially naïve households)

[89] Dan M. Kahan, Foreword, *Neutral Principles, Motivated Cognition, and Some Problems for Constitutional Law*, 125 HARV. L. REV. 1, 19–30 (2011); Philip E. Tetlock et al., *Social and Cognitive Strategies for Coping with Accountability: Conformity, Complexity, and Bolstering*, 57 J. PERSONALITY & SOC. PSYCH. 632, 632 (1989).

[90] BARON, *supra* note 52, at 208–11; George A. Akerlof & William T. Dickens, *The Economic Consequences of Cognitive Dissonance*, 72 AM. ECON. REV. 307, 307-09 (1982).

[91] Neil D. Weinstein & William M. Klein, *Resistance of Personal Risk Perceptions to Debiasing Interventions*, 14 HEALTH PSYCHOL. 132, 139 (1995).

[92] Suzanne C. Thompson et al., *Illusions of Control, Underestimations, and Accuracy: A Control Heuristic Explanation*, 123 PSYCH. BULLETIN 143, 144–61 (1998).

[93] Howard Kunreuther & Mark Pauly, *Rules Rather Than Discretion: Lessons from Hurricane Katrina*, 33 J . RISK & UNCERTAINTY 101, 105-06 (2006).

[94] *See* sources cited *supra* note 61.

A. The Information Problem

We saw in Part I.A. that regulators usually need two sets of information in order to optimally correct market failures.[95] First, they must know the marginal harm inflicted by each additional unit of the regulated good. For externality goods, this is the damage done to others, while for internalities it is the damage to the actor herself. At the same time, government needs to know how socially costly it will be to correct the problem: what is the actor's cost of keeping that next ton of carbon out of the air, that Twinkie out of our mouths?

Early criticisms of "paternalistic" regulation claimed that identifying the "harm" of an internality was inconsistent with basic economic methods,[96] but the literature has largely rejected that argument. That is, we usually infer marginal benefit from revealed preferences: the consumer faces a price, and if she is willing to pay that price, we conclude that her subjective welfare is greater than the price she paid.[97]

How, then, are we to second-guess consumers' choices? The answer is time.[98] By observing individual behavior over time, we can see whether people regret some of their own decisions, or take steps (commitments) to prevent themselves from making bad choices.[99] In this way, we can still rely on revealed preferences.[100] The concept of an "internality" doesn't necessarily privilege long-run over short-run preferences; we can treat them equally by simply adding them up, in effect balancing the revealed value of long-term preferences, such as regret and commitment, against the revealed value of momentary, System 1 preferences.[101] But since long-run preferences last much longer, they often will greatly outweigh those that last only fleetingly.[102]

Recent critics, such as Yale's Alan Schwartz, raise the more subtle problem that it can be challenging to identify which choices are bad decisions.[103] Suppose that the same action may be taken both by biased actors and the unbiased. Think of an auto insurance policy with a high

---

[95] *See supra* notes 29–45.

[96] Zamir, *supra* note 15, at 237–38 (attributing this view to John Stuart Mill).

[97] Paul Burrows, *Rationality and the Instrumentalist Case for Free Choice*, 15 INT'L REV. L & ECON. 489, 491 (1995); Daniel M. Haybron & Anna Alexandrova, *Paternalism in Economics*, *in* Coons & Weber, *supra* note 5, at 157, 159–60.

[98] ELSTER, *supra* note 2, at 5–7, 22; O'Donoghue & Rabin, *supra* note 25, at 1829 n.12.

[99] *Id.* ("[F]or any tax policy that takes effect in the future…the agent agrees that [the long-run perspective] is the appropriate welfare function.").

[100] *Id.*; Sunstein, *supra* note 5, at 1875–76; Zamir, *supra* note 15, at 247, 253. Although this point has been well-established for a decade, critics of internality correction continue to assert that any internality regulation unfairly privileges long- over short-term preferences. *E.g.*, Riccardo Rebonato, *A Critical Assessment of Libertarian Paternalism*, 37 J. CONSUMER POL'Y 357, 370 (2014).

[101] *Id.* at 246–47.

[102] Markus Haavio & Kaisa Kotakorpi, *The Political Economy of Sin Taxes*, 55 EUR. ECON. REV. 575, 578 (2011); B. Douglas Bernheim & Antonio Rangel, *Toward Choice-Theoretic Foundations for Behavioral Welfare Economics*, 97 AM. ECON. REV. 464, 467 (2007).

[103] PETER CSERNE, FREEDOM OF CONTRACT AND PATERNALISM 55 (2013); Hill, *supra* note 10, at 445–48; Schwartz, *supra* note 19, at 1377–78.

deductible.[104] That policy is a good choice for safe drivers, who don't need expensive coverage. It's a bad choice for bad but overconfident drivers: those who believe wrongly that they don't need a lot of coverage. Schwartz's claim is that, since we can usually only observe the consumer's choice to buy a particular policy, we can't know whether that choice produces internalities (the overconfident driver) or not.[105]

1. Solutions from the Externality Literature

The information problem is not unique to internality regulation. As we've seen, optimal regulation of an externality demands information about both the marginal damage or benefit produced by the good as well as the cost curves of private actors involved in producing it. Many key choices in modern theories of regulation turn on how best to reveal these necessary data.[106]

To be sure, not all the tools of externality regulation can be translated seamlessly to internalities. Price instruments serve as a favorite tool for revealing preferences and private costs,[107] but it is uncertain how well they operate for irrational actors. If government taxed high-deductible plans, that would screen out marginal over-confident drivers, but those that were especially over-confident might still make the same choice. That is, the change in price does not necessarily tell us whether the choice the consumer reveals is their "true" preference or a mistake.[108]

Time, another standard component of the externality tool kit, will often still perform well, however. Schwartz's credit contract examples are instances of so-called "ex ante" regulation—an effort to set policy and enforce it before consumers make their choice.[109] Many commentators argue that, when information is scarce, the better option is an "ex post" regulation.[110] That is, we wait until after the choice is made, measure the resulting harm or benefit, and apply a

---

[104] I owe this example to Goldin & Lawson, *supra* note 6, at 4.

[105] *Id.* A related claim sometimes raised in the anti-paternalism literature is that government regulators, too, can suffer from cognitive biases, or will be subject to market capture in ways that reduce their ability to process information. CSERNE, *supra* note 103, at 52, 54 (2013); Wright & Ginsburg, *supra* note 61, at 1063–64. These same difficulties arise in the standard externality setting, Zamir, *supra* note 15, at 280–81, and have standard solutions. For example, notice and comment rulemaking in its modern form helps to expose, and incentivizes agencies to limit, these kinds of failures. Mark Seidenfeld, *Cognitive Loafing, Social Conformity, and Judicial Review of Agency Rulemaking*, 87 CORNELL L. REV. 486, 508–46 (2002). For a more comprehensive empirical response to the bias and capture arguments, see Jeremy Blumenthal, *Expert Paternalism*, 64 FLA. L. REV. 721, 730–50 (2012).

[106] Helfand et al., *supra* note 42, at 251, 287; Steven Shavell, *Corrective Taxation Versus Liability as a Solution to the Problem of Harmful Externalities*, 54 J.L. & ECON. S249, S258–59 (2011).

[107] Kaplow & Shavell, *supra* note 40, at 4; Fullerton et al., *supra* note 39, at 430.

[108] *See* Galle, *supra* note 14, at 863–64 (discussing cognitive problems with drawing inferences from price instruments).

[109] Donald Wittman, *Prior Regulation vs. Post Liability: The Choice Between Input and Output Monitoring*, 6 J. LEGAL STUD. 93, 93 (1977)

[110] RICHARD A. POSNER, ECONOMICS ANALYSIS OF LAW 490–91 (8th ed. 2011); Jon D. Hanson & Kyle D. Logue, *The Cost of Cigarettes: The Economic Case for Ex Post Incentive-Based Regulation*, 107 YALE L.J. 1163, 1278 (1998).

corresponding price change. Torts are the classic example.[111] We can't easily predict which soda drinkers will be most prone to diabetes, and ban or limit their soda consumption.[112] But we can simply observe which consumers develop diabetes later, and allow them to sue to collect compensatory damages. [113]A rational, forward-looking beverage producer will anticipate the possibility of tort liability and act accordingly—although, as I have argued, limited liability and other factors may sometimes interfere with that process.[114]

We could take an analogous approach to the high-deductible example. For example, we might restrict high-deductible plans for individuals who have already demonstrated a history of risky driving, or even behaviors we know to be highly correlated with risky driving, such as regular alcohol consumption. More generally, we can observe what portion of the population that selects high deductible plans turns out to be relatively high risk. If the share of those who demonstrably made bad choices is high enough, it may still be welfare-improving on net to limit those plans, even if the plan would have improved well-being for some drivers.[115]

Admittedly, the expedient of switching from ex ante to ex post regulation may not work for all kinds of internalities. Many forms of ex post regulation, such as the tort suit, require the internality sufferer to recognize, at some point, that they have incurred harms.[116] This is plausible for Box 1 and Box 4 consumers—those who weren't attentive enough to notice their errors at the time, or lacked the willpower to avoid them. These individuals are likely to experience regret later. In contrast, Box 3 consumers may deny that their choices were wrong, and indeed may even harden their viewpoint further in order to avoid the cognitive dissonance that would come with acknowledging contrary evidence.[117]

In any event, externality theory also shows us that the informational demands of regulation for all kinds of internality sufferers can be much lower than Schwartz and other critics seem to assume. As William Baumol famously argued, sometimes all that we need to know is where we are relative to the optimal level of regulation.[118] Whatever might be the optimal number of tons of greenhouse gas a coal factory can emit, we know that the current level

---

[111] *Id.*

[112] Victor Fleischer, *Curb Your Enthusiasm for Pigouvian Taxes*, 68 VAND. L. REV. 1673, 1704–05 (2015).

[113] *Cf. id.* at 1705 (suggesting that cure for informational problem is to impose costs on obesity, not its predictors).

[114] Brian Galle, *In Praise of Ex Ante Regulation*, 68 VAND. L. REV. 1715, 1734–48 (2015).

[115] *See* Peter Diamond, *Consumption Externalities and Imperfect Corrective Pricing*, 4 BELL J. ECON. & MGMT. SCI. 526, 528–30 (1973) (deriving optimal pigouvian tax when externalities vary by consumer); *see* Zamir, *supra* note 15, at 266–67 (applying this principle to internalities).

[116] *See* William L.F. Felstiner et al., *The Emergence and Transformation of Disputes: Naming, Blaming, and Claiming*, 15 L. & SOC'Y REV. 631, 632–36 (1980–81).

[117] Eva Jones et al., *Confirmation Bias in Sequential Information Search: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information*, 80 J. PERSONALITY & SOC. PSYCH. 557, 557 (2001). Dan M. Kahan, *The Cognitively Illiberal State*, 60 STAN. L. REV. 115, 145–50 (2007), draws on primary psychological literature to argue that self-defensive bolstering can be minimized if new information is introduced within a frame that allows the listener to identify some elements that affirm her existing worldview.

[118] William J. Baumol, *On Taxation and the Control of Externalities*, 62 AM. ECON. REV. 307, 307–08 (1972).

authorized in the U.S. is too high.[119] We may not be able to identify the level at which there would no longer be additional social returns from further reduction. But we know we are far from that point and can safely lower emissions somewhat from the current, unregulated, level.[120]

Similarly, there are many internalities where we have at least this level of certainty. Again, economists estimate the internality cost of a pack of cigarettes at more than $30, for example.[121] Typically, the methodology is to measure the unbiased value of "good" outcomes in the general (and presumably unbiased) population, and assume that biased actors share that value. Thus Gruber and Koszegi, for instance, calculate the cost of smoking by looking at medical costs and the average person's value of additional years of life.[122] This assumes, of course, that smokers place equal value on long life.[123] Maybe that is a plausible assumption, but maybe it isn't. Still, even if smokers placed only half the value on life as others, Gruber's numbers at least tell us that current cigarette taxes are far too low.[124]

In my work on the choice between ex ante and ex post regulation, I also show that we can reduce the amount of information government needs in order to regulate, even ex ante, by using multiple prices or policies.[125] The social cost of a mistaken policy grows exponentially with the size of the mistake.[126] Through some simple mathematical simulations, I show that sorting actors into high, medium, and low risk categories can be just about as good as having perfect information about them.[127] Assuming government assigns a regulated party to the right category, the size of the error it's making—the distance, say, between the optimal price for that party and the actual price imposed—is smaller than if there were only one category, and the social cost accordingly declines exponentially.[128] We don't have to get policy choices exactly right in order for them to be good policies.

2. New Solutions
In addition to these familiar tools of externality regulation, there also are a host of new techniques, many still in development, aimed at the added informational problems raised by internalities. While these tactics may not cure every informational shortfall, they at least free

---

[119] *See* James R. Hines, Jr., *Taxing Consumption and Other Sins*, 21 J. ECON. PERSPECTIVES 49, 53, 64 (2007).

[120] *Id*. at 57; *see also* Louis Kaplow, *Optimal Control of Externalities in the Presence of Income Taxation*, 53 INT'L ECON. REV. 487, 488 (2012) (arguing that this proposition is always true "if a distribution-neutral income tax adjustment is employed" together with the externality correction).

[121] Gruber & Koszegi, *supra* note 71, at 1291.

[122] *Id*. at 1290–91.

[123] *Cf. Thomas Kneisner et al., Policy Relevant Heterogeneity in the Value of Statistical Life: New Evidence from Panel Data Quantile Regressions*, 40 J. RISK & UNCERTAINTY 15, 17 (2009) (reporting variations in estimates across income levels).

[124] Gruber & Koszegi, *supra* note 71, at 1292; JONATHAN GRUBER & BOTOND KÖSZEGI, A MODERN ECONOMIC VIEW OF TOBACCO TAXATION 17 (2008) (estimating internality-correcting price of about $14 per pack).

[125] Galle, *supra* note 114, at 1730–34.

[126] Kaplow & Shavell, *supra* note 40, at 775–79.

[127] Galle, *supra* note 114, at 1731–34.

[128] *Id*.

many internality problems from the most serious informational obstacles regulators might otherwise face.

The most familiar of these tools is asymmetric regulation.[129] Asymmetric regulations are more stringent for those who are most likely to make mistakes.[130] The now-classic example is default savings plans, under which employees must actively opt out of making retirement contributions.[131] Under-saving for retirement seems mostly to be caused by inattention and procrastination.[132] These are the same individuals who are the least likely to take the time to fill out the forms needed to opt out of the default savings plan.[133] Meanwhile, active savers who prefer lower or different retirement savings will readily fill out the one-page form, and so bear little cost from the default.[134] Unless even active savers are making mistakes, it follows that asymmetric instruments will always be preferable to an outright mandate.[135] Plan designs that force individuals to decide whether to opt in or out can also be asymmetric; for those for whom decisions are not burdensome, the cost of deliberation is trivial.[136]

Because asymmetric regulation is self-targeting, regulators don't have to be able to identify biased consumers.[137] The regulation applies to everyone. Since the cost of non-compliance is so minor for fully rational actors, though, it doesn't change their behavior.[138] Therefore regulators can apply it to the whole population without worrying much about the risk that it will be misapplied.[139] This same analysis still largely holds if the regulation is more than just a minor inconvenience for the rational.[140] As I explained, when government can set multiple

---

[129] Colin Camerer et al., *Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism,"* 151 U. PA. L. REV. 1211, 1230–37 (2003).

[130] *Id*. at 1222, 1225–26; Hunt Allcott et al., *Energy Policy with Externalities and Internalities*, 112 J. PUB. ECON. 72, 74–75 (2014).

[131] John Beshears et al., *The Importance of Default Options for Retirement Savings Outcomes: Evidence from the United States*, *in* SOCIAL SECURITY POLICY IN A CHANGING ENVIRONMENT 167, 187–92 (Jeffrey Brown et al. eds., 2009).

[132] *Id*. at 183–84.

[133] *Id*. at 188.

[134] *Id*.; *see* Chetty et al., *supra* note 78, at (reporting that "active savers" were highly responsive to changes in incentive savings, implying low transaction costs). Bubb and Pildes emphasize that there may be more than one dimension of heterogeneity among a regulated population. Bubb & Pildes, *supra* note 6, at 1621–26. The group of passive savers may include individuals who would have (eventually) saved more than the default savings rate, such that the default actually lowers savings for those individuals. *Id*. at 1622. In my view this is a problem created largely by the use of only a single instrument. If government is limited to only one instrument, the default, then there are tradeoffs implicit in the setting of the level of default savings. A more efficient approach would be to use a second instrument to further sort the inattentive savers between those with high and low savings needs. Allcott et al., *supra* note 130, at 77–78; *cf.* Ayres, *supra* note 22, at 2093 (noting possible use of multiple defaults for "subsets" of population).

[135] Goldin & Lawson, *supra* note 6, at 3.

[136] *Cf.* Beshears et al., *supra* note , at 188 (advocating "active choice" savings in some settings).

[137] Camerer et al., *supra* note 129, at 1222.

[138] *Id*.

[139] *Id*. at 1225–26.

[140] *See* Ayres, *supra* note 22, at 2089–91.

prices or policies, the social cost of imperfectly-informed regulation is much lower.[141] Asymmetric regulations are essentially just regulations with multiple, built-in prices: a high price for irrational actors, a lower price for the fully rational.[142]

Other regulatory techniques work by inducing consumers to reveal their potential biases, enabling regulators then to target the right policy to the right person. A simple example of this "separating equilibrium" approach, discovered accidentally by tax officials, is over-withholding.[143] Millions of taxpayers each year voluntarily allow their employers to withhold more in taxes each pay period than required, or opt to receive their tax refund in one lump sum rather than incrementally over the year.[144] Both qualitative and quantitative studies show that taxpayers are, in effect, using the government as a commitment device, forcing themselves to save until the time of their tax rebate.[145] In general, we should expect that these kinds of opt-in regulations will be attractive to individuals in Box 4: those who recognize their need for, and who value, interventions to bolster willpower.[146]

We could probably tell a similar story about Box 1, the inattentive. Many of us know that we sometimes fail to pay as much attention as we should to some of life's important

---

[141] The relationship between the size of the government's errors and social welfare, and the possibility of multiple prices, also undermines Schwartz's suggestion that law should default to an assumption that actors are rational. Schwartz, *supra* note 19, at 1403–04. In effect, Schwartz is proposing that we should set the pigouvian price on internalities to zero, unless we have compelling evidence otherwise. But that is a disastrous policy, because it greatly increases the average expected distance between the government's price and the optimal. Say that there is a 50% chance that the optimal price is $100. This is no different, statistically, than saying that half the population has an optimal price of $0 and half $100. In that case, optimal price should be set at $50. Allcott & Sunstein, *supra* note 6, at 17; *see* Allcott et al., *supra* note 130, at 76 (modeling argument that optimal internality tax is always above zero if any consumer is biased, assuming biases are in one direction). Even better, if the government has any information about which consumers are more likely to need a $100 correction, it should impose a $100 tax on those consumers, and a $0 tax on others.

[142] Galle, *supra* note 114, at 1754. Avishalom Tor, *The Next Generation of Behavioral Law & Economics*, *in* EUROPEAN PERSPECTIVES ON BEHAVIOURAL LAW AND ECONOMICS 17, 25–26 (Klaus Mathis ed. 2015), offers a similar story. Tor explains that legal rules can themselves cause "selection" effects, such as when risky rules tend to cause the affected population to be more risk-seeking.

[143] Barr & Dokko, *supra* note 59, at 979; Richard H. Thaler, *Anomalies: Saving, Fungibility, and Mental Accounts*, 4 J. ECON. PERSP. 193, 193-95 (1990).

[144] Damon Jones, *Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds*, 4 AM. ECON. J.: POL'Y 158, 158 (2012).

[145] Barr & Dokko, *supra* note 59, at 979; Sara Sternberg Greene, *The Broken Safety Net: A Study of Earned Income Tax Credit Recipients and a Proposal for Repair*, 88 NYU L. Rev. 515, 561–62 (2013); Jones, *supra* note 144, at 159; Damon Jones, *Information, Preferences, and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401(k) Savings*, 2 AM. EC. J.: APPLIED EC. 147, 149 (2010) (reporting reasons the advanced EITC program failed); Ruby Mendenhall et al., *The Role of Earned Income Tax Credit in the Budgets of Low-Income Earners*, 86 SOC. SERV. REV. 367, 377–78, 382, 398 (2012).

[146] Nava Ashraf et al., *Tying Odysseus to the Mast: Evidence from a Commitment Savings Device in the Phillipines*, 121 Q.J. ECON. 635 (2006); Esther Duflo et al., *Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya*, 101 AM. ECON. REV. 2350 (2011); *see also* Ashvin Gandhi & Michael Kuehlwein, *Reexamining Income Tax Overwithholding as a Response to Uncertainty*, 43 PUB FIN. REV. __, manuscript at 3 (forthcoming 2016) (reporting evidence that rules out most plausible rational-actor explanation for overwithholding).

details.[147] Here, too, there are robust commercial markets for self-help devices, providing evidence that indeed some individuals will value interventions.[148] Our slenderer readers may be unfamiliar with Weight Watchers.[149] The Weight Watchers "points" system is basically just an easily-implemented tool for encouraging participants to pay attention to what they eat and drink, and to count calories and other nutritional data.[150]

A third set of tools, as Schwartz at points acknowledges, is experimentation and data crunching.[151] Consider, for instance, the possibility that some consumers make bad choices because of features of their choice environment, such as in studies finding that long menus of Medicare Part D drug coverage options caused seniors to pick plans that were clearly dominated by other available choices.[152] Schwartz seems to assume that these situations are hopeless, at least in the case where every choice would be rational for some consumers.[153]

As Goldin and Reck recently have shown, however, government can design experiments that at least would reveal what share of the population's choice has been changed by the framing of the choice.[154] "Consistent" consumers, whose choice is unaffected by the frame, presumably are choosing based on some set of invarying, underlying preferences, and therefore their expressed preferences are reliable evidence of their true preferences.[155] Government can examine which observable features of consistent choosers predict a given set of preferences, and then use the same observables to draw inferences about the preferences of inconsistent choosers whose preferences can't be directly measured.[156] As I have pointed out elsewhere, these same types of tools can also be used to extend experiments, especially field experiments, of limited scope to conclusions about the population as a whole.[157] While it is well known that some

---

[147] See Raj Chetty et al., *Salience and Taxation: Theory and Evidence*, 99 AM. ECON. REV. 1145, 1170–74 (2009) (modeling behavior of households aware of their own inattention).

[148] Michael S. Barr et al., *Behaviorally Informed Regulation*, in BEHAVIORAL FOUNDATIONS OF PUBLIC POLICY 440 (Eldar Shafir ed., 2013). A notable data point here is the continuing popularity of software that allows us to plan and reminds us of those plans. A number of studies, summarized in Brigitte Madrian, *Applying Insights from Behavioral Economics to Policy Design*, 6 ANN. REV. ECON. 663, § 3 (2014), find that planning and reminders improve savings, vaccination rates, education performance, vehicle safety, and healthy eating.

[149] https://www.weightwatchers.com/us/

[150] https://www.weightwatchers.com/us/our-approach, last visited Jan. 21, 2016 ("Our new SmartPoints™ plan nudges you toward a healthier pattern of eating so that over time, smart choices become second nature").

[151] Schwartz, *supra* note 19, at 1380, 1402–03.

[152] Jason Abaluck & Jonathan Gruber, *Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program*, 101 AM. ECON. REV. 1180, 1195–96, 1198 (2011); Saurabh Bhargava et al., Chose to Lose? Employee Health-Plan Decisions from a Menu with Dominated Options, unpublished manuscript, Nov. 2014, at 21–22.

[153] Schwartz, *supra* note 19, at 1390–93, 1403 n.54.

[154] Jacob Goldin & Daniel Reck, Preference Identification Under Inconsistent Choice, unpublished manuscript, Mar. 25, 2015, at 7–29. A less technical version of the same argument is Jacob Goldin, *Which Way to Nudge? Uncovering Preferences in the Behavioral Age*, 125 YALE L.J. 226, 260–69 (2015).

[155] Goldin & Reck, *supra* note 154, at 4–5.

[156] *Id*. at 5–6; *see also* Allcott et al., *supra* note 130, at 78–79 (suggesting that government can at least infer bounds on extent of bias if consumers are fully responsive to prices).

[157] Galle, *supra* note 14, at 862–64.

cognitive failings are context-specific or overlapping,[158] this is not a reason to reject experimentalism, but rather a reason to design experiments so that they will have external validity.[159]

Ian Ayres and Quinn Curtis offer the kernel of a similar idea in their recent proposal to reform pension savings.[160] They suggest that government simply test directly for financial sophistication before it allows investors to choose from outside the limited default set of investment options.[161] It won't always be practical or cost-effective to administer individualized testing, especially for decisions that must be made quickly or frequently. But the direct-testing idea can be implemented for other infrequent, high-stakes choices, such as home mortgages, student loans, and health plans.

B. Government or Markets?

Jonathan Klick & Gregory Mitchell and others also argue against internality regulation on the ground that it simply crowds out what would otherwise have been private responses.[162] They point out that government regulation of internalities may reduce individual incentives to invest in the power to self-regulate, which they call "cognitive hazard."[163]

It isn't clear the cognitive hazard story makes much sense for Box 3 internalities. Confirmation-biased actors by definition resist the notion that they are biased at all.[164] This may also hold for some Box 1 individuals: inattentive actors may well lack awareness of their own inattention.[165] In both these cases individuals would not invest much effort in de-biasing themselves, whether the government helped or not.

In other instances, as we have seen, there are already private markets for internality correction, but government intervention can be justified even when these private markets function efficiently.[166] One rationale for regulation is a public taste for distributive justice.[167]

---

[158] Schwartz, *supra* note 19, at 1392–95.

[159] *Cf.* Sendhil Mullainathan et al., *A Reduced-Form Approach to Behavioral Public Finance*, 4 ANN. REV. ECON. 17.1, 17.3 (2012) (laying out a model of cognitive failures that "can be interpreted independent of a specific psychological mechanism").

[160] Ian Ayres & Quinn Curtis, *Beyond Diversification: The Pervasive Problem of Excessive Fees and"Dominated Funds" in 401(k) Plans*, 124 YALE L.J. 1476, 1525–29 (2015).

[161] *Id*. at 528.

[162] Jonathan Klick & Gregory Mitchell, *Government Regulation of Irrationality: Moral and Cognitive Hazards*, 90 MINN. L. REV. 1620, 1626, 1633–41 (2006); CSERNE, *supra* note 103, at 53; Gary Becker & Casey Mulligan, *The Endogenous Determination of Time Preferences*, 112 Q.J. ECON. 729 (1997); Epstein, *supra* note 60, at 811–17.

[163] Klick & Mitchell, *supra* note 162, at 1633–41; *see also* GARY S. BECKER, ACCOUNTING FOR TASTES 11 (1996) (arguing that individuals can invest in long-run thinking); Schwartz, *supra* note 19, at 1379 (arguing that government protection is unnecessary when consumers can learn from their own mistakes).

[164] Kahan, *supra* note 89, at 22–23.

[165] *Cf.* O'Donoghue & Rabin, *supra* note 25, at 1841 (making this point about willpower failures).

[166] To the contrary, in my view there are good reasons to believe markets for learning are frequently imperfect, whether due to limited opportunities for learnings, Oren Bar-Gill & Elizabeth Warren, *Making Credit Safer*, 157 U. PA. L. REV. 1, 7– 25(2008), or other market imperfections, Brian Galle, *Carrots, Sticks, and Salience*, 67 TAX L.

We may believe that individuals who have the bad luck to suffer from internalities should not be left worse off than others.[168] Many commentators offer similar theories for providing affordable health insurance regardless of economic need: women and individuals with disabilities, for instance, should not be worse off economically than men, simply because their health costs are on average higher.[169]

Obliging internality sufferers to turn to private-market solutions will often leave them worse off because private help is not free. Self-education is difficult and often impractical.[170] In many cases, private market solutions may allow the solution provider to extract all or nearly all of the welfare surplus from customers.[171] For example, employers who provide pensions help workers to overcome the workers' failure to save adequately for retirement. In addition to sacrificing some salary for this benefit, workers also take on the agency costs inherent in entrusting their bargaining adversary–the employer–with their own well-being.[172]

Even worse, private markets may extract consumer surplus without necessarily overcoming consumers' internalities. Another lesson from studies of tax withholding is that after households engage in costly commitment devices, they later backslide and, when tempted, spend more money trying to accelerate the very payments they voluntarily delayed.[173] These de-commitment services, known as "refund anticipation loans" in the tax refund context, are often subject to enormous fees.[174] Thus, as other commentators have observed, the possibility that

---

REV. 53, 102–03 (2013); *see* Jeremy A. Blumenthal, *Emotional Paternalism*, 35 FLA. ST. UNIV. L. REV. 1, 51–53 (2007). My argument here holds whether these suppositions are ultimately empirically supported or not.

[167] Concerns for justice are consistent with my welfarist framework, since welfarism aims to maximize the satisfaction of all preferences, including society's preferences for fairness. MATTHEW D. ADLER & ERIC A. POSNER, NEW FOUNDATIONS OF COST-BENEFIT ANALYSIS 23 (2006); Louis Kaplow & Steven Shavell, *Fairness vs. Welfare*, 114 HARV. L. REV. 961, 989–92 (2001).

[168] SHLOMI SEGALL, HEALTH, LUCK, JUSTICE 100–04 (2010); Tom Baker & Peter Siegelman, *You Want Insurance With That? Using Behavioral Economics to Protect Consumers from Add-on Insurance Products*, 20 CONN. INS. L.J. 1, 43 (2013); *cf.* DANIELS, *supra* note 62, at 161–62 (suggesting that inequality of information about health risks can warrant regulation of those risks).

[169] SEGALL, *supra* note 168, at 74–86, 101; Eric Rakowski, *Who Should Pay for Bad Genes?*, 90 CAL. L. REV. 945, 1352–67 (2002); *see also* Zamir, *supra* note 15, at 282–83 (offering this rationale in defense of paternalistic regulation more generally).

[170] Bar-Gill & Warren, *supra* note 166, at 26–55, and Galle, *supra* note 166, at 104–05, survey the evidence.

[171] *See* Galle & Utset, *supra* note 83, at 54–55 (describing this phenomenon in the consumer credit context).

[172] Christine Jolls, Employment Law and the Labor Market, NBER Working Paper No. 13230, at 4 (2007); Brendan Maher, *Regulating Employment-Based Anything*, 99 MINN. L. REV. __, manuscript at 43–48 (forthcoming 2015).

[173] *See* Barr & Dokko, *supra* note 59, at 980; Jonathan Parker, *Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment*, NBER Working Paper No. 21369 (July 2015) (reporting that borrowing to accelerate refunds is most strongly correlated with measures of household impatience).

[174] Barr & Dokko, *supra* note 59, at 979; Chi Chi Wu & Jean Ann Fox, Nat'l Consumer Law Ctr., Inc., Consumer Fed'n of Am., Coming Down: Fewer Refund Anticipation Loans, Lower Prices from Some Providers, but Quickie Tax Refund Loans Still Burden the Working Poor 4, 8–12 (2008).

private markets can both offer *and undo* self-commitments can be a reason for government action.[175]

Even if private providers do not extract most of the benefits of internality correction, government provision of commitment devices and the like can be an efficient source of redistribution because it resembles an improved version of the income tax. In standard public finance accounts of the tax system, government can improve over a simple tax on labor earnings if it can identify and tax things that are correlated with the ability to earn income.[176] Briefly, the reasoning is that in a progressive tax system, individuals with high earning potential may "mimic" low-income individuals in order to escape high tax rates; taxing correlates of income rather than income itself makes this mimicking more difficult.[177]

Free internality correction fits into this story. The same logic behind taxing correlates of income also justifies government provision of in-kind benefits, if those benefits are disproportionately useful to individuals with lower earning potential.[178] Many of the key building blocks of internalities, such as impatience, inattention, and addictive behaviors, have been shown to correlate with lower income.[179] Internality correction therefor closely resembles an efficient supplement to the income tax.

Even if none of these redistributive rationales hold, it may still make sense for government to protect consumers from wasteful investments in private internality reduction. Many internalities are the result of manipulations by other actors who can profit from consumer mistakes.[180] In those cases, consumer efforts to resist manipulation are akin to investments in theft prevention: like window-bars and RFID-resistant wallets, cognitive "hardening" forces

---

[175] David Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 Q.J. ECON. 443 (1997); O'Donoghue & Rabin, *supra* note 25, at 1841.

[176] This insight is usually traced to James A. Mirrlees, *Optimal Tax Theory: A Synthesis*, 6 J. PUB. ECON. 327 (1976), and the earlier work by Mirrlees discussed therein.

[177] Emanuel Saez, *The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes*, 83 J. PUB. ECON. 217, 226, 228 (2002).

[178] LOUIS KAPLOW, THE THEORY OF TAXATION AND PUBLIC ECONOMICS 227 n.10; Helmuth Cremer & Phillipe Pestieau, *Redistributive Taxation and Social Insurance*, 3 INT'L TAX & PUB. FIN. 281, 282 (1996); Jean-Charles Rochet, *Incentives, Redistribution, and Social Insurance*, 16 GENEVA PAPERS ON RISK & INSURANCE THEORY 143, 160–64 (1991).

[179] For evidence on impatience, see James J. Heckman et al., *The Rate of Return to the Highscope Perry Preschool Program*, 94 J. PUB. ECON. 114 (2010); Walter Mischel et al., *Delay of Gratification in Children*, 244 SCIENCE 933 (1989); Brian C. Cadena & Benjamin J. Keys, Human Capital and the Lifetime Costs of Impatience, unpublished manuscript, May 2014, *available at* http://spot.colorado.edu/~cadenab/Research_files/cadena_keys_lifetime_impatience.pdf. For salience, see Jacob Goldin & Tatiana Homonoff, *Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity*, 5 AM. ECON. J.: ECON. POL'Y 302 (2013). For addiction, see, e.g., Karen M. Jennison, *The Short-Term Effects and Unintended Long-Term Consequences of Binge Drinking in College: A 10-Year Follow-Up Study*, 30 AM. J. DRUG & ALCOHOL ABUSE 659 (2009), or watch a few episodes of *The Wire*.

[180] Willis, *supra* note 6, at 1170–73.

consumers to reallocate resources they could have spent on something they actually wanted.[181] Regulation protects consumers against this waste.[182]

Finally, Klick & Mitchell do not consider the possibility that government support could encourage private self-help investments. Government support might, for example, increase the returns to investment: I might invest more effort in learning to rely on commitment devices if the devices available are cheap and effective.[183] Absent government assistance, some individuals might be too demoralized to make use of commitment or procrastinate learning how to do so.[184] In essence, regulation serves as a commitment matching grant, multiplying the public's investments.

IV. Choice of Instruments: Carrot, Stick, or Compromise?

Assuming that officials have committed to regulating internalities, how should they do it? Should we ban dangerous drugs? Tax them? Subsidize alternatives and treatments? Or oblige producers to print visceral and frightening images of the drugs' consequences on the side of each package? These choices among regulatory options or "instruments" is a central question for modern regulation.[185]

I have argued before that there are four basic categories of regulatory instrument.[186] Two of the categories involve explicit transfers of wealth, while the other two don't.[187] If government selects an instrument that transfers wealth, it must decide whether the regulated party will, relative to the pre-existing baseline, be made to pay for non-compliance (a "stick") or be rewarded for compliance (a "carrot").[188] The transferless categories include standard "command and control" regulation, such as bans or caps on production.[189] Many nudges resemble command and control regulation, in that they also do not involve any explicit wealth transfer.[190] But other

---

[181] Steven Shavell, *Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 COLUM. L. REV. 1232, 1239 n.28 (1985).

[182] *Id.*; *see also* Baker & Siegelman, *supra* note 168, at 44–45 (pointing out that seller efforts to create consumer mistakes are also economically wasteful).

[183] *Cf.* Gruber & Koszegi, *supra* note 71, at 1284 (arguing that knowledge of a future price increase may help motivate addicts to quit addictive behavior before price increase).

[184] *See* Galle & Utset, *supra* note 83, at 76–78 (introducing the concept of "meta-procrastination" and discussing its sources and implications).

[185] Helfand et al., *supra* note 30, at 249–53.

[186] Galle, *supra* note 14, at 848–54. I actually simplify my earlier categories a bit here for the sake of narrative economy. A more complete categorization would also include the division between ex ante and ex post, and would sub-divide price instruments into priced (those denominated in dollars) and transfer (those that shift resources). Transfer instruments can be even further refined to distinguish between public and private transfers.

[187] *Id.*

[188] *Id.* at 851.

[189] *Id.* at 848–49.

[190] *Id.* at 846–47.

features of the nudge are so different and, as Sunstein and Thaler say, "surprising," from a classical rational-actor perspective,[191] that I put them in a fourth category by themselves.

As we saw briefly in Part I.A., there now is a considerable literature examining the factors that officials should consider when they choose between these classes of instrument. Instruments may differ, among other ways, in their propensity to create moral hazard, in their effects on demand for the regulated product, in their impact on the public fisc, and in their distributive fairness.[192] While economists often favor wealth-transfer instruments over others,[193] the choice between carrot and stick typically obliges us to trade off favorable results on some of these factors against unfavorable outcomes on others, implying that sometimes a reason to choose a transferless instrument is that it may represent a compromise position.[194]

Although prior literature has extensively considered these factors in the externality context, it has not to my knowledge examined the extent to which the received wisdom applies to internalities. This Part begins that task.[195]

A. Moral Hazard

Moral hazard is usually the strongest argument against carrots.[196] When government offers rewards to polluters to stop their misdeeds, it encourages new polluters to begin emitting so that they, too, can be paid to stop.[197] The underlying game theory logic is similar to the rationale of the United States government when it refuses to pay ransom to kidnapers.[198] In contrast, the promise of future liability encourages polluters to invest in pollution-reduction technologies, especially if liability may turn out to be retroactive;[199] rewards provide no incentive to remedy harm before the reward is paid, and may crowd out any voluntary compliance.[200] Even when distributional concerns might weigh in favor of carrots,[201] the

---

[191] THALER & SUNSTEIN, *supra* note 4, at 85–86, 252–54.

[192] Galle, *supra* note 14, at 848–53.

[193] Helfand et al., *supra* note 30, at 287; Edward L. Glaeser, *Paternalism and Psychology*, 73 U. CHI. L. REV. 133, 150 (2006); Kaplow & Shavell, *supra* note 40, at *passim*.

[194] Galle, *supra* note 14, at 872–79.

[195] I emphasize "begin." A work of this length necessarily must omit many considerations, some potentially important. For example, the form of regulation may affect compliance and the public's relation to the law. Yuval Feldman & Orly Lobel, *Behavioral Tradeoffs: Beyond the Land of Nudges*, San Diego Legal Studies Paper No. 14-15B, manuscript at 12–28, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2439179. Private actors who profit from consumer may respond to government interventions, and understanding how these responses might vary across instruments is an important future direction for researchers. *Cf*. Ran Spiegler, *On the Equilibrium Effects of Nuding*, J. LEGAL STUDIES __ (forthcoming 2016) (offering formal models of firm responses to several nudges).

[196] Ronald H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 42 (1960); Wiener, *supra* note 21, at 726 & n.186.

[197] *Id*. For qualifications to this story, see Galle, *supra* note 21, at 822–23.

[198] *See* ROBERT MNOOKIN, BARGAINING WITH THE DEVIL: WHEN TO NEGOTIATE, WHEN TO FIGHT 7 (2010).

[199] Saul Levmore, *Changes, Anticipations, and Reparations*, 99 COLUM. L. REV. 157, 1663 (1993).

[200] WILLIAM J. BAUMOL & WALLACE E. OATES, THE THEORY OF ENVIRONMENTAL POLICY 212 (2d ed. 1988) (1975).

[201] Gerrit De Geest & Giuseppe Dari-Mattiacci, *The Rise of Carrots and the Decline of Sticks*, 80 U. CHI. L. REV. 341, 363–65 (2013).

possibility of these kinds of moral hazard strongly counsels in favor of at least partial stick liability for externality producers.[202]

This version of the moral hazard argument seems tenuous when it comes to internalities. A rational actor can't credibly kidnap themselves.[203] Nor would it make sense to delay self-help in order to encourage government payment, unless somehow the individual expects that the government will greatly over-pay. Although Klick & Mitchell argue that internality-regulation creates a form of moral hazard, discouraging individuals from investing in private de-biasing,[204] that would seem equally true no matter what form the government's policy takes. At worst, there may be some differences in timing. A newly-introduced carrot might encourage "sophisticated" high-beta consumers to wait to kick their bad habit until the government assistance arrives, while a new tax would encourage quitting before the tax.[205]

For now, though, the possible absence of significant moral hazard concerns is enormously important to internality regulation. With moral hazard out of the way, the choice between carrot and stick is much closer.[206] Carrots for internalities deserve serious consideration.[207] Let's continue a bit further in that direction.

B. Income Effects

As other writers recognize, income effects can be important in the choice of regulatory instrument.[208] Income effects are the just the tendency of most consumers to demand more of a good when they have more money. A relatively unusual exception is the so-called "inferior" good, which is a good that people tend to want less when their wallet is fatter; ramen noodles may be a familiar example to college-educated readers. Proponents of soda taxes note that

---

[202] In essence, a partial stick is similar to a government co-pay requirement on its social insurance policy, and is efficient for similar reasons: it trades off the worst of moral hazard against only a small loss on redistribution. Raj Chetty & Amy Finkelstein, *Social Insurance: Connecting Theory to Data*, *in* V HANDBOOK OF PUBLIC ECONOMICS 112, 157–58 (Alan J. Auerbach et al. eds., 2012).

[203] But see *Other People's Money* (1991).

[204] *See supra* notes 162–165.

[205] *Cf.* Gruber & Koszegi, *supra* note 71, at 1273–77 (finding that smokers are responsive to anticipated future tax increases). It also is possible that some interventions may be more likely to cause crowding out of private effort than others. For example, maybe resentment towards fines imposed on helmetless riders tends to strengthen confirmation bias among motorcycle riders, making them even more resistant to news about the dangers of riding. This is another area where better empirical evidence would be useful.

[206] Galle, *supra* note 14, at 878.

[207] My analysis of carrots here is limited to what I'll call traditional carrots, those that have no unexpected impacts on decisions. It may be possible to use rewards as the centerpiece of a "surprising" regulatory intervention, in which the behavioral impact is far larger than we would expect given the size of the reward. See Kevin Volpp et al., *Financial Incentive-Based Approaches for Weight Loss*, 300 JAMA 2631, 2631--36 (2008) (describing effectiveness of constant stream of very small but immediate rewards in changing the behavior of present-biased actors). Since the fiscal costs of these approaches is typically quite small, I will group them with nudges and other transferless instruments.

[208] THOMAS STERNER, POLICY INSTRUMENTS FOR ENVIRONMENTAL AND NATURAL RESOURCE MANAGEMENT 167 (2003).

because of income effects, subsidies for healthy alternatives could have a perverse impact.[209] By enriching a household, subsidies may also increase its demand for all consumables, including unhealthy food.

More formally, we could say that with subsidies the private marginal cost curve shifts leftwards, reflecting the fact that for any given level of subsidy, the amount of junk food foregone is less than we would have expected absent the income effect.[210] Assuming that the marginal internal benefits curve is downward sloping, this shift implies a higher optimal subsidy (or, putting the point more intuitively, when consumers are more reluctant to give up their junk food, government suasion must be more powerful). This higher subsidy in turn drives some additional income effect, and so on. Figure three illustrates.

Figure Three: Income Effects and Optimal Price
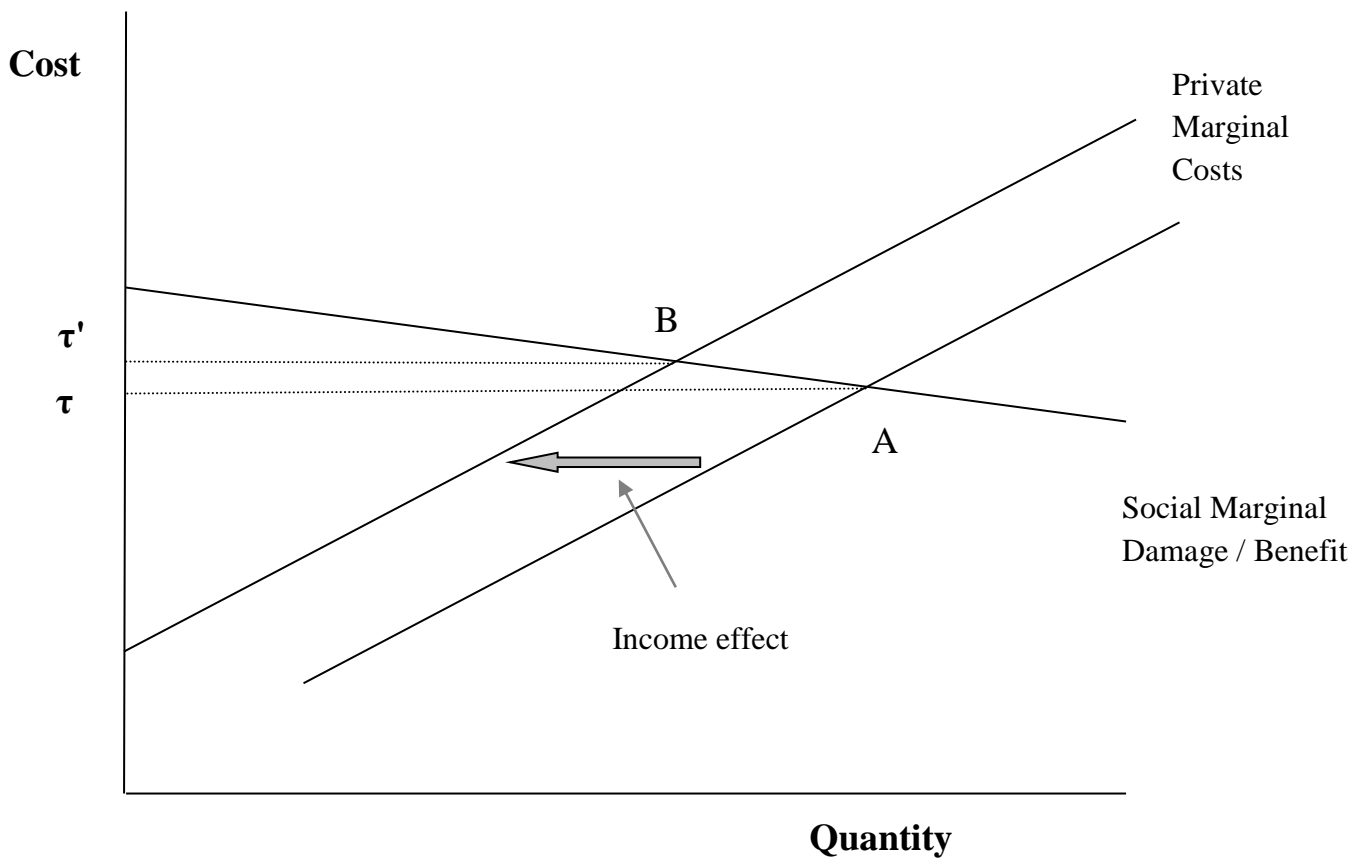
[209] Gio   *Fig. 3: Income effects of a subsidy shift private costs curve left, raising optimal price*   *'e*
Thin S
[210] *See* KAPLOW, *supra* note 120, at 496–97 (analyzing impact of income effects on optimal pigouvian tax).

The opposite is true if government uses sticks rather than carrots, or if government uses carrots but the internality-creating item is an inferior good. In those cases, income effects will depress consumption, allowing for a lower optimal tax or subsidy rate.

Regulation of internalities may also change a household's perception of its own wealth, but this turns out likely not to affect the choice of instruments analysis. If we can achieve the same internality correction in any of three ways (carrot, stick, or transferless instrument), presumably any income effect that comes with internality correction will be the same under each.[211] For example, helping families save for retirement might make them feel richer; since consumption is a normal good, that feeling of greater wealth might increase current expenditures at the cost of savings.[212] But this effect would be equally true whether we penalize non-savers or reward savers. It is only the incremental income effects that result from the carrot/stick choice that factor into which of those options we would want.

Of course, that wouldn't be true if our choice of instrument also affected how individuals perceive the benefit of internality-correction. For instance, we saw before that inattentive eaters might not be aware that a "nudge" has changed their eating habits. They might be somewhat more cognizant of monetary rewards, leading them to better understand that they now are making better use of their money (and therefore are effectively richer). I explore the relation between salience and income effects in more detail elsewhere.[213]

Ultimately, then, the importance of income effects is hard to predict without more empirical evidence. There is some suggestion already that "junk" food and cigarettes might be inferior goods, at least in the contemporary United States.[214] If other internality-creating choices are also inferior goods, then the argument for carrots becomes much stronger. In that case, the income effect of the carrot boosts, rather than undermining, the substitution effect of the price instrument. If not, alternatives such as sticks or transferless instruments look better by comparison, unless it were the case that carrots had particularly low income-salience.

---

[211] *See* Louis Kaplow, *Myopia and the Effects of Social Security and Capital Taxation on Labor Supply*, NBER Working Paper No. 12452, at 7 (Aug. 2006), for a discussion of the potential income effects of improving a household's inter-temporal budget allocation.

[212] *See* GRUBER, *supra* note 30, at 650 (stating that income effects tend to reduce savings because present consumption is a normal good).

[213] Galle, *supra* note 166, at 66–67, 86–89, 93.

[214] R.J. DeGrandpre et al., *Effects of Income on Drug Choice in Humans*, 59 J. EXPERIMENTAL ANAL. BEHAVIOR 483, 483 (1993) (generic cigarettes); Leonard H. Epstein et al., *Purchases of Food in Youth: Influences of Price and Income*, 17 PSYCH. SCIENCE 82, 88 (2006) (junk food); Matthew Harding & Michael Lovenheim, The Effect of Product and Nutrient-Specific Taxes on Shopping Behavior and Nutrition: Evidence from Scanner Data 16 (Mar. 4, 2013) (unpublished manuscript) (junk food).

C. Revenues

Another central issue in the choice of instruments is the instrument's effect on the public fisc. Carrots, of course, must be paid for.[215] Carrots aside, in most of the prior literature, the fact that some instruments provide funds that can be transferred to others in large measure explains the dominance of "price," or what I have called "transfer," instruments, over other choices.[216]

O'Donoghue & Rabin suggest that internalities present an especially strong example, going so far as to argue that Pigouvian-type taxes on internalities can be justified solely from a revenue perspective.[217] This was an argument that, in the pollution context, had been known in the 1990's as the "revenue recycling" or "double dividend" claim: the possibility that carbon taxes both could clean the environment and also raise money more efficiently than other sources.[218] Allcott & Sunstein assert, without reference to O'Donoghue & Rabin, that there is a double dividend from taxing internalities, but as we'll see the story is rather more complicated.[219]

The availability of a double dividend is, fittingly, important to my analysis for two reasons. For one, a double dividend would offer a reason to impose a tax on internalities, regardless of whether we otherwise would believe internality regulation is a good idea: sin taxes would just be a particularly good way to raise money. The other significance of the double dividend is that it represents an opportunity cost if policy makers choose to use command and control regulation or nudges, rather than taxes or other transfer instruments.

1. Is There a Double Dividend?

To understand O'Donoghue and Rabin, we first have to review briefly why it is that environmental economists came to mostly reject the "double dividend" theory. In brief, as Bovenberg & Goulder summarize in their handbook entry, the problem is that carbon taxes are differentiated consumption taxes.[220] As a result, carbon taxes are likely not only to distort consumers' choice of what to buy, but also to affect the after-tax return on labor income. That is, since we generally work in order to buy stuff, taxes on the stuff itself affects not only our choice about what to buy but also our labor/leisure decisions.[221] While carbon taxes may have some desirable effects, they also frustrate some consumers' preferences for high-carbon goods, and discourage labor supply.[222]

[215] Madrian, *supra* note 148, at § 2.

[216] *See* sources cited *supra* note 193.

[217] O'Donoghue & Rabin, *supra* note 25, at 1829, 1832--33.

[218] *E.g*., Charles L. Ballard & Steven G. Medema, *The Marginal Efficiency Effects of Taxes and Subsidies in the Presence of Externalities: A Computational General Equilibrium Approach*, 62 J. PUB. ECON. 199, 200 (1993).

[219] Allcott & Sunstein, *supra* note 6, at 8.

[220] A. Lans Bovenberg & Lawrence H. Goulder, *Environmental Taxation and Regulation*, *in* 3 HANDBOOK OF PUBLIC ECONOMICS 1471, 1486–1507 (Alan J. Auerbach & Martin Feldstein eds., 2002)

[221] Anthony B. Atkinson & Joseph E. Stiglitz, *The Design of Tax Structure: Direct Versus Indirect Taxation*, 6 J. PUB. ECON. 55, 56 (1976)

[222] Bovenberg & Goulder, *supra* note 220, at [21–22].

Boiling Bovenberg & Goulder's more complex calculations down into a simple equation, the effect of a carbon tax is:

$$E - C - L + R$$

where E is the gains from carbon reduction, C the loss from consumers' choice to switch to a second-best product, L the compensated labor-supply impact, and R any available gains from cutting other distorting taxes.[223]

This equation implies that pollution taxes are usually a less efficient revenue source than the income tax. Set aside E for now; these environmental gains can be achieved through non-tax policies. Bovenberg and Goulder argue that L and R will at best cancel each other out, because the carbon tax, like the income tax, discourages labor supply.[224] Since L and R each involve identical dollar amounts, the result is that the carbon tax is strictly worse than the income tax by the amount of the distortion, C.[225]

O'Donoghue and Rabin in essence point out that this result does not hold for internalities when taxes are too small to create any effects on labor supply.[226] By definition, correction of the internality on net improves consumer welfare, at least from the point of view of the government planner.[227] If there are negligible labor-supply effects, we could write the resulting equation as:

$$U = I - C + R$$

where I replaces E as gains from the government policy, and by assumption $I - C$ is always positive. That seems quite intuitive; O'Donoghue and Rabin's main contribution is to show that

---

[223] We could also write a slightly more complicated version of the formula, with the added complexity perhaps justified by the fact that this version captures an important nuance. This more-nuanced version would look something like:

$$E - C/p_C^2 - L/p_L^2 + R/p_R^2$$

Here, C and L stand for something a bit different: instead of the total welfare loss, they are the total loss, in dollars, caused by the tax's distortive effects. These losses are spread out over a population, p, which differs for each factor depending on the incidence of that factor. To translate the dollar loss into a utility loss, we divide the dollars by the square of the population that experiences the dollar loss. This reflects, roughly, the basic proposition in tax economics that the deadweight loss of a tax rises in proportion to the square of its rate: small taxes don't inconvenience us much, but large taxes motivate ever-larger distortions in our behavior. By spreading a burden out over more people, we reduce its per-capita impact, effectively cutting the rate.

This burden-spreading point is important to Bovenberg & Goulder. For them, it helps to explain why R can never offset C and L. By bringing in carbon-tax revenue, the government can cut the income tax by the same amount, producing a beneficial effect R/p_R^2. But the income tax affects many more people than will be impacted by at least some forms of carbon pricing. R and L are equal, by definition. If p_R is much greater than p_L, the R/p_R^2 term will always be smaller than L/p_L^2.

[224] Bovenberg & Goulder, *supra* note 220, at [22].

[225] *Id*. at [23–24].

[226] O'Donoghue & Rabin, *supra* note 25, at 1834--35.

[227] *Id*. at 1829; *see also* Haavio & Kotakorpi, *supra* note 102, at 575.

it sometimes is plausible one could achieve meaningful changes in behavior with taxes that are too small (in their total burden, not necessarily their rate) to affect labor/leisure decisions.[228]

But what about when internality-correcting taxes must be large in this sense? Cigarette taxes, for example, can consume a meaningful fraction of household income for low-income families.[229] In many cases, it is not convincing to ignore potential labor-supply (or other tax-avoidance) effects.

A 2006 article by Louis Kaplow provides some starting points for our analysis.[230] Kaplow analyzes the labor-supply effects of an actuarially fair social security system.[231] He shows that the labor-supply effects of compelled savings for retirement may depend on taxpayer beliefs about whether their own savings decisions are optimal.[232] Let us consider that point in the context of the four boxes I set out in Part II.

Of the internality examples I have examined so far, Box 3 of Fig. 2 most closely resembles the standard externality case sketched by Bovenberg & Goulder. These are individuals, again, who perceive change as costly now and see little long-term benefit to compliance. Think of a tobacco tax. Individuals who change their tobacco consumption habits are likely to experience that shift as very costly. Further, individuals who believe that they do not face major long-term costs of smoking will perceive themselves as obtaining a lower utility for each dollar of earnings, potentially diminishing labor supply or shifting labor to less-productive uses. At the same time, there are revenue gains from the tax, and the policy planner recognizes that there is in fact a long-run internality health gain for the individual, yielding the simple equation:

---

[228] O'Donoghue & Rabin, *supra* note 25, at 1836–38. Another contribution the pair offer is that internality-correcting taxes can be efficient even if the population is heterogeneous, such that taxes in some cases are falling on individuals with no internality at all. The intuition is that taxes on rational actors, if "small," are relatively non-distorting, because at the margin the rational actor is indifferent between her two options; if taxes change her choice, the welfare loss is correspondingly small. That is, the contribution of rational actors to C in my formula is minor. At the same time, because internality sufferers are far from their private optimum, small changes can produce large welfare gains (due to the exponential relationship between deadweight loss and the size of the error). Thus, I will generally be quite large relative to C. *See also* Mullainathan et al., *supra* note 159, at 17.16–17.17 for a slightly more elaborate model of the same idea.

[229] *See* Jonathan Gruber & Botond Köszegi, *Tax Incidence When Individuals Are Time-Inconsistent: The Case of Cigarette Excise Taxes*, 88 J. PUB. ECON. 1959, 1962 (2004) (estimating that smoking consumes about 3% of the budget of households in the bottom quartile of income); Gary Lucas, *Saving Smokers from Themselves: The Paternalistic Use of Cigarette Taxes*, 80 U. CIN. L. REV. 693, 693 (2012) (estimating that a New York pack-a-day smoker pays $2,500 per year).

[230] Kaplow, *supra* note 211. Portions of the 2006 paper were published as Louis Kapow, *Targeted Savings and Labor Supply*, 18 INT'L TAX & PUB. FIN. 507 (2011). As the working paper version includes much more extensive discussion of several subjects of interest here, citations are to the original.

[231] *Id*. at 4–13. Daniel Shaviro, *Multiple Myopias, Multiple Selves, and the Undersaving Problem*, NYU Law & Economics Working Paper No. 382 (Aug. 2014), available at http://lsr.nellco.org/cgi/viewcontent.cgi?article=1386&context=nyu_lewp, considers a similar case but relaxes the assumption of actuarial fairness.

[232] Kaplow, *supra* note 230, at 12–13.

$$U = I \ - C - L + R \qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(Eq. 1)}$$

This of course is the same as Bovenberg & Goulder's, with internalities replacing externalities.

Now consider Box 4, where individuals have high $\hat{\alpha}$ and also high $\hat{\beta}$. Imagine, for instance, the reaction to an alcohol tax by an actress we could call "Lindsay," who recognizes her alcohol dependence but lacks the willpower to end it.[233] Lindsay would experience the loss of her morning whiskey as a drop in her short-run wellbeing, even while she recognizes that there are long-run gains from abstaining. To the extent she does not climb fully onto the sobriety wagon, government would have revenue from taxing alcohol.

What is the effect on Lindsay's labor supply? As in Kaplow's story of the social security tax, arguably Lindsay should recognize that the government, by improving the way in which she has chosen to allocate her spending, has on net increased her returns to labor.[234] This effect should at least partly counteract the labor-discouraging impact of the tax, and might even on net increase her labor.[235]

At the same time, there are income effects that point in the opposite direction from the substitution effects. For example, since Lindsay is poorer as a result of the liquor excise, she will have to work harder in order to pay her rent. Conventionally, analyses of the labor impact of taxation omit this effect, because by assumption the consumer is "compensated" by the use of the tax to pay for public goods, which offset the loss of individual wealth.[236] But there is no such compensation for another kind of income effect Lindsay experiences. When the government helps her to set her priorities in order, she in effect is wealthier: by incentivizing her to spend less money on booze, government gives her more money to pay rent.[237] Since she now has more money available to pay rent, the government's assistance should tend to diminish her labor supply.

By adding subscripts to the labor component to reflect the differing impact of the two competing substitution effects and the income effect, we could write the resulting social welfare as:

$$U = I - C \ + L_{s1} - L_{s2} - L_i + R \qquad\qquad\qquad\qquad\qquad\text{(Eq. 2)}$$

Equation 2 implies that there is some possibility of a double dividend from taxing internalities of this kind. Even setting aside I, it is possible that the remaining terms net out to a

---

[233] Any resemblance of our anecdote to any real individuals, living or dead, is strictly coincidental.

[234] Kaplow, *supra* note 230, at 7.

[235] Lindsay may also be able to work more as a result of her greater sobriety. Some portion of that improvement should be reflected in the *I* term in our equation. There may also be externality benefits for others from Lindsay's efforts. But for simplicity I will focus on the pure internality case.

[236] GRUBER, *supra* note 12, at Ch. 21.

[237] Kaplow, *supra* note 230, at 9; *cf.* Chetty et al., *supra* note 147, at 1173–74 (exploring possible income effects of improved allocation of resources for irrational consumers).

positive, implying that there would be social gains from using the Pigouvian tax to replace other sources of revenue.[238] Whether this is so in the real world would depend on the relative labor gains and losses from better allocating consumers' budgets.

Box 1 offers an even stronger case for double dividends. Recall that these individuals tend not to notice either changes in their behavior or to contemplate the long-run effects of those changes. What is the effect of a tax on inattentive behaviors?

Changes to the behavior of the inattentive may have minimal effects on either consumer welfare or labor supply.[239] If I have little idea how much soda I drank, it is unlikely that I will perceive changes in that amount as affecting my short-run well-being.[240] And, since I am unaware that anything meaningful has changed, I have no reason to adjust my labor supply. Of course, it is also quite possible that incentives that depend on conscious responses, such as tax or tax incentives, also will not do much to change my behavior, producing little internality gain,[241] but that would make them potentially a very effective source of revenue.[242] In the best-case scenario, assuming minimal consumption losses or labor effects, but modest internality gains (for instance, because consumers are more attentive to price than they are to portion size) we could write the welfare effects of the tax on inattentive consumers as:

$$U = I + R \qquad\qquad\qquad\qquad\qquad\qquad\text{(Eq. 3)}$$

which would represent an unambiguous double dividend.[243]

On the other hand, taxes aimed at the highly inattentive could also be highly inefficient, depending on how they are designed and how taxpayers respond. For example, repealing existing tax incentives for retirement savings might be perceived by inattentive savers as an

---

[238] Given the possibility of private "commitment devices" for individuals in Box 4, the social planner may prefer to calculate I as the incremental gains, if any, from public provision.

[239] Galle, *supra* note 14, at 867–68.

[240] *See* ELSTER, *supra* note 2, at 8 (noting that inattentive agents cannot respond to incentives); Kahneman, *supra* note 50, at 1451 ("The operations of System 1 . . . are difficult to control or modify").

[241] O'Donoghue & Rabin, *supra* note 25, at 1835; Shaviro, *supra* note 231, at 34–43. For a survey of the evidence on taxpayer responsiveness to "low-salience" taxes, see Galle, *supra* note 166, at 63–67; David Gamage & Darien Shanske, *Three Essays on Tax Salience: Market Salience and Political Salience*, 65 TAX L. REV. 19, 26–53 (2011).

[242] Usually, the welfare loss from a "hidden" tax of this kind is that the consumer obtains a different basket of goods than she would have purchased had she been aware of the tax. Chetty et al., *supra* note 147, at 1173. By assumption here, the consumer does not have clearly defined preferences for the amount of the internality good. But the tax would reduce her consumption of other goods, to an extent she likely would not have chosen if she were aware of the tax. Consumers may be able to minimize these distortions if they notice their shrinking budget and can re-optimize purchases accordingly. *Id.* at 1174. So while labor distortions are lessened, there may still be welfare losses from this misallocation of the taxpayer's budget. If so, we should add a $-C$ term, which would make the case for a double dividend ambiguous.

[243] Kaplow, *supra* note 230, at 8 & n.10, appears to tell a similar possible story about willpower failures. He suggests that, if workers wrongly believe at the time they set labor effort that the government will not impose any commitments on them, that there will be no labor supply effects. But, unlike the inattention story I describe, there is no known empirical evidence of this behavior. And, in a more realistic setting in which labor supply is constantly being decided, it is unclear how long workers could persist in their mistaken beliefs about the government's response. At some point, Kaplow's willpower story becomes an inattention story.

increase in the tax on labor.[244] A tax on non-savers could be similarly ineffective, if inattentive savers are attentive to labor taxes. The penalty would distort labor income as much as any other income tax, but if the non-savers don't know why they're being punished it may not change behavior in the desired direction. More research on these issues would be welcome.

In sum, whether an internality-correcting tax provides a double dividend depends on the nature of the mental processes that affect consumer choice. Further complicating the analysis, there may be a mix of processes that produce similar outcomes. Some smokers may fall into Box 3, others in Box 4; that is, some may want help to quit, while others believe they don't need to "right now."

Optimal policy choice then may depend on government's ability to correctly identify the mix of each. This task may not be as daunting as it sounds. For example, we expect that individuals in Box 4 will accept commitment devices where available (barring, say, ideological objections to accepting any help from the government), while individuals in the other boxes would not. This difference allows government to establish policies that induce individuals in Box 4 to reveal their type.

2. Double dividends and choice of instruments

Assuming that in some cases there may be a double dividend, let's look at how that possibility would affect the government's choice of instruments. Here again, it will make a big difference which "box" our regulated party falls into. Once more, prior commenters have argued that the transfer of resources from the regulated party to other private individuals or the public makes price instruments unambiguously superior, on welfare terms, to transferless regulation.[245] That is, since the welfare effect of a carbon tax is $E - C - L + R$, while the effect of a similar command & control regime is presumably $E - C - L$, the tax is always superior to the extent of any revenue gains.[246]

The same could be true of internality regulation. To the extent that consumers are aware of the effect of the transferless regulation, we might expect the regulation to provide the same labor-supply effects that a tax would. That is, if the regulation is experienced as an unwanted psychic cost (the hassle of going back for a second tiny cup of soda, say), it will diminish the pleasure of the drinking experience, lowering the returns to labor.[247] If government policies are perceived as helpful because they lead to better allocation of the consumer's spending choices, transferless regulation should match the impact of a tax or other "stick," but lack the corresponding revenue.

In earlier work, I attempted to show that this syllogism is not true to the extent that government has available transferless instruments that have lesser effects on consumer welfare

---

[244] Kaplow, *supra* note 230, at 1–2.
[245] *See* sources cited *supra* note 193.
[246] *Id.*
[247] Glaeser, *supra* note 193, at 150.

and labor supply than the carbon tax.[248] For example, nudges and other "surprising" instruments may sometimes change consumer behavior without consumers necessarily noticing that much important has changed.[249] If so, then it becomes ambiguous which instrument is better on basic utility terms, as $E - C - L + R <> E - C$.

To be more concrete, consider the choice between raising tobacco taxes and the global efforts to label cigarette packaging with disturbing images of adverse health outcomes.[250] Let's posit that some fraction of smokers fall in Box 4 of Figure 1; that is, they would prefer to quit but lack the willpower to do so, and appreciate government efforts to motivate their cessation efforts.

It is unclear if the images are inferior on revenue terms. Both options, if effective, would encourage greater labor effort by improving the smoker's perceived returns to working (because she smokes less, improving her health), while reducing incentives to work because the smoker will perceive herself to be richer. The higher tax, in addition, brings in revenue, and as we just saw may even on net produce a double dividend. The disturbing images, of course, do not bring in money (and may reduce revenue if the government maintains a tobacco excise at a lower rate).

Whether nudges are the better choice than taxes turns on the relative labor effects of the nudge and the tax. Do disturbing images make smokers feel less inclined to work, because the discomfort they feel when they smoke reduces the total reward they can buy with their labor effort? If so, then taxes are superior: both choices have similar labor-supply effects, while taxes bring in money. But if not, then even if there is a double dividend from cigarette taxes the nudge is likely a better choice. Taxes have lower labor and more revenue, while the nudge has less revenue but more labor supplied.[251] That is, the nudge is superior if

$$(I - C + L_{s1} - L_{s2} - L_i + R < I - C + L_{s1} - L_i) = (R < L_{s2}) \qquad \text{(Eq. 4)}$$

I showed in my earlier work that under standard assumptions $R < L_{s2}$: the nudge is preferable to the tax.[252]

That is also the case, *a fortiori*, for Box 3 taxpayers. As we saw, there is no double-dividend scenario for taxing individuals in Box 3.[253] If nudges are better even when using them means giving up a double dividend, they must be better when it doesn't.[254]

Perhaps surprisingly, nudges may lose out for Box 1 individuals, the attentive. In the best-case scenario for taxation I just sketched, inattentive actors do not adjust labor supply in

---

[248] Galle, *supra* note 14, at 867–71.

[249] *Id*. at 867–68.

[250] *See* sources cited *supra* note 8.

[251] Note that since labor is taxed, the extra labor supply under a nudge also has a revenue benefit. For simplicity, I simply assume that this benefit is already reflected in the L terms.

[252] Galle, *supra* note 14, at 869–71.

[253] *See supra* notes 232–233.

[254] In terms of my simple equations, a nudge is better for Box 3 taxpayers when $I - C - L + R < I - C - L$, which is to say never.

response to changes in their consumption. Switching to a nudge then sacrifices the double dividend, without gaining any off-setting labor supply benefits. In math terms, $I + R > I$.

In sum, it looks preliminarily as though "nudges" and other surprising instruments are the best choice for consumers who experience willpower or bolstering problems, while sin taxes make more sense for the inattentive. Again, we don't yet fully understand the labor-supply effects of many internality-correction options. If the labor supply effects of sin taxes for the inattentive fall short of the best-case scenario, that might tip the balance back towards nudging.

D. Information and Targeting

Another standard argument in favor of price instruments over regulation is that they provide better information about private costs.[255] Typically government cannot directly observe the private marginal cost of compliance.[256] However, a price instrument induces rational externality producers to comply if their private costs are less than the price set.[257] By iterating this process, government can gather enough data about private cost structures to better approximate the optimal price.[258] This information is also potentially critical to effective targeting of a policy: government does not want to distort the behavior of those who are already performing in their own self-interest.[259]

Of course, for our purposes the critical assumption in the standard account is that the observed response to price is a rational one. Yet we already know that in many cases it is not. Eighty-five percent of Danish workers ignored a large new tax incentive for retirement savings.[260] Does that tell us that the cost of retirement savings was very high, or just that Danes prefer not to think about Act Five of their lives? It seems that in many cases price instruments

---

[255] Don Fullerton et al., *Environmental Taxes*, *in* DIMENSIONS OF TAX DESIGN: THE MIRRLEES REVIEW 423, 430 (James Mirrlees et al. eds., 2010). Kaplow & Shavell, *supra* note 40, at 4.

[256] *Id*.

[257] *Id*.

[258] Jeff Strnad, *Conceptualizing the "Fat Tax": The Role of Food Taxes in Developed Economies*, 78 S. CAL. L. REV. 1221, 1254–55 (2005).

[259] *Id*.; *cf.* Kaplow, *supra* note 230, at 22 (discussing significance of policy choice when population is heterogeneous in their propensity for error); Tor, *supra* note 142, at 26–27 (same).

    Allcott & Sunstein rely on a version of the targeting argument to favor energy subsidies over clean-fuel mandates, but their analysis may be a bit off. They suggest that a subsidy will be superior to a "mandate that all consumers take action" because the mandate will cause compliance among some consumers for whom compliance is inefficient. Allcott & Sunstein, *supra* note 6, at 7. This confuses the form of an instrument with its price. Allcott & Sunstein appear to assume that the mandate would apply to every consumer–in effect, that its price would be infinite. But transferless instruments, including many command & control approaches, can have an effective or "shadow" price of any amount. Galle, *supra* note 14, at 862. To take the simplest example, the mandate could simply exempt any consumer with private compliance costs above what would have been the subsidy amount. This would effectively set the price of compliance at *tau* for either instrument. The duo acknowledge this point later in their discussion, Allcott & Sunstein, *supra* note 6, at 9, so perhaps we should understand their claim simply to be that a flat ban on inefficient energy use by all consumers is bad policy, rather than a general claim about the merits of taxes over mandates.

[260] Chetty et al., *supra* note 71, at [ ].

are no better than others at revealing private information.[261] We have seen that Schwartz and others rely on this fact as a basis for objecting to any form of internality regulation.

I argued earlier that experiments and self-targeted instruments answer Schwartz's critique, and they also serve to level the playing field between price and other instruments. Because the Danish experiment had a control group of individuals whose costs of savings were indistinguishable from the treatment group—and that control group responded strongly to the tax incentive—we can infer that the unresponsiveness of the bulk of the population was due to behavioral factors, not cost.[262] Well-designed experiments like this allow government not only to identify individuals who need a little help, but also to draw inferences about the private cost structures of the targeted groups.[263] In a regulatory environment in which government is already conducting experiments before it regulates, the need to rely on price instruments to reveal information is considerably lessened.

If anything, price instruments might be less appealing in an internality context because they may be more difficult to design as self-targeting. "Linear" taxes, or taxes that apply a uniform per-unit rate, are difficult to make asymmetrical.[264] It's true, of course, that a teetotaler won't pay much alcohol tax.[265] But the large man who *rationally* consumes alcohol at a slow, steady pace will pay far more than the slender woman who irrationally binges. That is poor targeting.

Nonetheless, with some creativity policy makers can likely reduce the inflexibility of tax-like instruments. Consider a system of opt-in taxation.[266] Individuals could agree to be subject to a higher tax rate on some goods. Ian Ayres's StickK, a company that allows users to pledge to pay a penalty if they fail to meet personal goals, has already adopted this method.[267]

Similarly, government could allow households to opt out of government subsidies for overly tempting products. Some states are currently considering a prohibition on junk food for families who receive SNAP benefits,[268] but a more empowering option that would also reveal better information would be an opt-in system in which households have the power to move

---

[261] *See* Mullainathan et al., *supra* note 159, at 17.8–17.9, 17.22.

[262] Chetty et al., *supra* note 71, at [ ].

[263] Galle, *supra* note 14, at 861–63.

[264] Fleischer, *supra* note , at 1686, 1697–1701; Haavio & Kotakorpi, *supra* note 102, at 576.

[265] O'Donoghue & Rabin, *supra* note 25, at 1831 (claiming that tax distortions on rational actors are "second order").

[266] Pratt, *supra* note 68, at 131–32. See Ted O'Donoghue & Matthew Rabin, *Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes*, 93 AM. ECON. REV. (Papers & Proceedings) 186, 190 (2003), for an early version of a voluntary sin tax proposal.

[267] IAN AYRES, CARROTS AND STICKS: UNLOCKING THE POWER OF INCENTIVES TO GET THINGS DONE Ch. 8 (2010).

[268] Anemona Hartocollis, *New York Asks to Bar Use of Food Stamps to Buy Sodas*, NEW YORK TIMES, Oct. 6, 2010, at A1; Steve Mistler, *Maine DHHS Renews Push for Ban on Buying Soda and Candy with Food Stamps*, PORTLAND PRESS HERALD, Nov. 23, 2015.

selected categories of food and beverage on or off a "banned" list.[269] Perhaps modifying the list could not be done in-store, which would help to reduce the likelihood that the family would buy the items it doesn't want to be tempted by. Manuel Utset and I have also proposed a kindred idea in the consumer credit context in which recipients of government rebates are defaulted into saving a portion, but have the power to access the funds in emergencies and to change the default savings level.[270]

What about the inattentive? Two Finnish economists, Haavio and Kotakorpi, argue that taxes cannot be made to vary with an individual's propensity to make mistakes, both because that is unobservable (they say) and because mistake-free buyers would purchase tax-free and then resell at a discount to the error-prone. [271] Transaction costs would often remove the second concern. Even assuming that both concerns were in full force, regulators could design around them by taxing observable behaviors that are correlated with internalities, not the internality itself.[272]

An example here could be a tax that scales up with portion size. The bigger the bottle of soda, or the more cigarettes per package, the greater the tax rate per ounce or per cigarette. Analyses of inattentive behaviors show that portion size is a major driver of consumption—or, to put it another way, the minor nuisance of having to acquire another serving slows consumption considerably among those who tend to be overconsumers.[273] Of course, retailers would respond to the scaled-up tax rate by selling smaller portions, but that is exactly the goal of the policy; by reducing portion size, we also reduce internalities. For rational actors, the bother of buying a second cup is trivial. We therefore have an instrument that, despite being a tax, is nonetheless asymmetric: rational actors won't pay it.

E. Distribution

Distributive considerations can also be important to choosing an instrument, and typically will favor transferless instruments.[274] Some are relatively straightforward. Critics of sin taxes

---

[269] *Cf.* Janet Schwartz et al., *Healthier by Precommitment*, 25 PSYCHOLOGICAL SCI. 538, 538–46 (2013) (reporting experiment in which consumers could trigger loss of an existing subsidy if they failed to improve their healthy shopping).

[270] Galle & Utset, *supra* note 83, at 84–87.

[271] Haavio & Kotakorpi, *supra* note 102, at 587.

[272] *Cf.* Mullainathan et al., *supra* note 159, at 17.18 & n.17 (suggesting that effective prices can be varied with degree of internality by concentrating enforcement based on observable correlates of internality).

[273] *See* Chandon, *supra* note 77, at 16 (connecting portion size and inattention to overeating); Andrew B. Geier et al., *Unit Bias: A New Heuristic That Helps Explain the Effect of Portion Size on Food Intake*, 17 PSYCHOL. SCI. 521, 5224 (2006) (suggesting that "immediate presence" of tempting goods drive the effects of portion size on consumption).

[274] While it is possible, and perhaps even preferable, for policy makers to analyze each rule as though it were enacted simultaneously with a perfectly offsetting adjustment to the income tax, Kaplow, *supra* note 120, at 488, 494, in practice this step may not always be feasible, *id*. at 499. Among other reasons, it may be difficult to observe the distributive effects of some policies, such as the self-targeting instruments I have described here. In that situation, Kaplow suggests analyzing the policy as though it were transferless, but immediately followed by an income tax adjustment with the appropriate redistributive characteristics. *Id*. at 498.

often complain that they fall more heavily on the poor.[275] Gruber and Koszegi respond that, by assumption, an internality-correcting tax on net improves well-being for each individual, so that if the tax falls more on the poor this simply means that it provides even greater benefits for the poor than its burdens.[276] But this latter analysis overlooks opportunity costs. Assuming that the government has available some other, transferless, instrument that potentially provides individuals with the same benefits as the tax, the tax will necessarily be more regressive than the tax-free alternative, unless it is offset by refunds or other support.[277]

Free or subsidized support for internality-sufferers can also be an efficient way of redistributing to the poor. In general, the distributive features of externality correction policies don't offer reasons to enact them, as the same redistributive benefits could be obtained through a simpler, less-distortive income tax.[278] Recall, however, that in-kind transfers can be more efficient than cash transfers via the tax system in the special case where the in-kind benefit disproportionately benefits individuals with low earning potential.[279] Then the government can support indigent households without creating incentives for individuals with high earning potential to stay home. Internality correction, we saw earlier, can have this feature. Since internality-correcting taxes or other "sticks" would undercut the benefit of efficient redistribution by diminishing the size of the net benefit the poor receive, these options are less appealing than others in some cases.

F. Summary

Let's try to pull together some of these threads. First, in many cases carrots are likely to remain the least favored policy. Carrots for internalities lack the ruinous moral hazard problems they pose for externalities, but they remain socially costly because they must be paid for through some kind of tax revenue. As a form of price instrument, carrots for externalities can generate better information than transferless policies, but we've seen that these information benefits will often be minor or redundant in the internality context. Thus, the extra cost of carrots is only likely to be worth paying if their income or redistributive benefits are large enough to justify the markup. Both of these factors can arise sometimes but are not at all universal.

Next, nudges and other "surprising" transferless instruments, such as defaults and choice architecture, look like potentially strong options for overcoming willpower or confirmation-bias failures, especially when willpower failures are correlated with low earning capacity. Again, there will often be little information lost by abandoning transfer instruments for nudges. The

---

[275] Lucas, *supra* note 229, at 738–39; Pratt, *supra* note 68, at 121–24.

[276] Jonathan Gruber & Botond Koszegi, *Tax Incidence When Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes*, 88 J. PUB. ECON. 1959, 1960 (2004); *see also* Avigail Kifer, The Incidence of a Soda Tax, in Pennies and Pounds, unpublished manuscript. Nov. 25, 2014, at 5 (making same claim about soda taxes).

[277] *See* Emmanuel Farhi & Xavier Gabaix, Optimal Taxation with Behavioral Agents, unpublished manuscript, Aug. 26, 2015, at 26; *cf.* Haavio & Kotakorpi, *supra* note 102, at 576 (noting that sin taxes transfer wealth from the irrational to the rational). On the refund possibility, see Hines, *supra* note 119, at 65.

[278] KAPLOW, *supra* note 178, at 32.

[279] *See supra* text accompanying notes 176–179.

nudge is less regressive than a tax would be, and may even serve as an efficient mode of supporting the poor. Depending on the labor-supply effects of a given nudge, it may be superior on revenue terms to a tax alternative. And even the income effects of the nudge are preferable in the case of tempting inferior goods.

The case is less clear cut for the inattentive. For these individuals, a tax potentially could be a very efficient revenue-raiser, although there would be serious distributive fairness concerns if it ended up falling mostly on lower-income households. The income effects of the tax are also more likely to be useful in most cases, since most consumption goods are ordinary.

Gruber & Koszegi distinguish between different forms of error in another way. They observe that naïve addicts may have to face a higher price, since they do not recognize the benefit of the government's policy.[280] If so, that also has implications for our choice of instruments, as higher prices in the form of taxes imply greater revenues but also great distributive effects.

It's worth emphasizing, too, that government needn't rely on a single instrument for any given social problem. As the externality literature recognizes, there are some good arguments for relying on multiple instruments, including my recent argument that it allows a way of imposing multiple ex ante price points.[281] In addition, we have seen that different cognitive failings may produce the same mistaken behavior. If some over-eaters are inattentive while others suffer willpower failures, it could make sense to use a different instrument to help each of the sub-populations.[282] In some cases, though, instruments may be in tension with one another; a nudge aimed at Box 4 might still reduce demand among Box 1 individuals, diminishing tax revenues earned from taxing the inattentive.[283] If these revenues were the main reason for adopting the tax, that combination might not make sense.

Another argument for multiple instruments arises if actors vary in their sensitivity to dollar instruments. In that case, I've shown, the optimal approach is to increase the price of the instrument, but not by so much as to fully correct the behavior of the most insensitive.[284] The intuition for this result is similar to the argument for using several ex ante prices: because the social cost of errors grows exponentially, it is better to make a few small mistakes than one big

[280] Gruber & Koszegi, *supra* note 71, at 1289–90.

[281] Galle, *supra* note 114, at 1730–33; Michael P. Vandenbergh et al., *Regulation in the Behavioral Era*, 95 MINN. L. REV. 715, 719 (2011); *see* Emanuel Saez, *The Optimal Tax Treatment of Tax Expenditures*, 88 J. PUB. ECON. 2657, 2659–60 (2004) (explaining use of tax subsidies and adjustments to level of direct government provision as complementary tools).

[282] Allcott et al., *supra* note 130, at 77–78. Similarly, if there are both consumption externalities and internalities, different instruments may be necessary to target both effectively. Madrian, *supra* note 148, at § 2; Goldin & Lawson, *supra* note 6, at 4.

[283] *Cf.* Kaplow, *supra* note 230, at 22 (noting that mandatory savings and savings subsidies produce inefficient results when combined).

[284] Galle, *supra* note 166, at 77–81.

one.[285] Even better, though, would be to eliminate one of the small mistakes. A second, behavioral instrument aimed at the group that is most insensitive to a traditional tax or subsidy would improve over the dollar instrument alone.[286]

Obviously, there remains an enormous amount of uncertainty with all of these prescriptions. My goal isn't to be able to provide definitive answers to, say, how to regulate vaccinations. The point instead is to identify for further study the factors that we need to know in order to make the best policy.

V. Application: Tobacco Regulation

To repeat, at this point there remain important unknowns in evaluating the best policy for any given internality. To make my analysis concrete, though, I offer some preliminary thoughts, given the available evidence, on how my argument would apply to a significant real-world source of internalities: smoking.

Recent U.S. efforts to follow other countries around the world in requiring that packages of cigarettes prominently display vivid images of smoking's health consequences were stymied by a panel of the U.S. Court of Appeals for the D.C. Circuit.[287] The panel ruled that the images infringed on the First-Amendment right of manufacturers to control their brand message, the administration declined to seek *certiorari*, and the government went back to the drawing board.[288] But two years later, the full D.C. Circuit, sitting en banc in a different dispute, ruled that the relatively searching scrutiny it had used in the earlier case was not justified in the commercial speech context.[289] Therefore it appears there is again an opportunity to revive the vivid-images rule.[290] What can the government say in defense of vivid images?

Smoking is a cognitively complex behavior, with different smokers seeming to exhibit different kinds of at least arguably irrational behavior. Evidence suggests some smokers are classic examples of Box 4 willpower failure.[291] Another contributing cause for some smokers is a form of inattention bias, in which the immediate visceral temptation of habitual cues that trigger the urge to smoke bypass rational thought processes.[292] These we might place in Box 1. A small group of smokers seem to reject evidence that their personal risks of smoking are

---

[285] *Id*. at 78–79.

[286] *See* Farhi & Gabaix, *supra* note 277, at 25, 28 (modeling combination of tax and nudge when some actors are insensitive to taxes).

[287] *R.J. Reynolds Tobacco Co. v. Food & Drug Admin*., 696 F.3d 1205, 1208 (D.C. Cir. 2012).

[288] Letter from Eric Holder, Att'y Gen., to John Boehner, Speaker, U.S. House of Representatives (Mar. 15, 2013), available at http://www.justice.gov/sites/default/files/oip /legacy/2014/07/23/03-15-2013.pdf.

[289] *Am. Meat Inst. v. U.S. Dep't of Agric*., 760 F.3d 18, 22–23 (D.C. Cir. 2014) (en banc).

[290] For more detailed discussion of the First-Amendment issues, see Rebecca Tushnet, *More Than a Feeling: Emotion and the First Amendment*, 127 HARV. L. REV. 2392, 2404–15, 2442–33 (2014).

[291] Gruber & Koszegi, *supra* note 71, at 1278; Joni Hersch, *Smoking Restrictions as a Self-Control Mechanism*, 31 J. RISK & UNCERTAINTY 5, 6 (2005).

[292] Bernheim & Rangel, *supra* note 47, at 44–45; George Loewenstein, *A Visceral Account of Addiction*, *in* GETTING HOOKED: RATIONALITY AND ADDICTION 235, 237–45 (Jon Elster & Ole-Jórgen Skog eds., 1999)

serious, although this group tends to be among the heaviest smokers.[293] That, of course, accords well with our Box 3.

To meet the challenge of this very heterogeneous group of consumers, tobacco control policy likely must be equally complex. For each sub-population of smokers, a different regulatory instrument may be optimal. Further, it may be optimal to use multiple instruments even within each sub-group. A number of studies show that vivid images have helped to motivate and encourage quitting and its follow-through, and to discourage adolescent smoking.[294] Which instrument or instruments, then, should we choose?

Let's take Box 4 willpower-failure sufferers to start. For them, it is fairly straightforward that the vivid images policy is the best choice, given one key factual assumption. If it is the case that the images don't have significant impacts on labor/leisure decisions or related distortions typically associated with a sales tax, then the images are on net preferable strictly on revenue terms. It's worth emphasizing we don't currently have evidence on that question, although it seems intuitive that those who have a long-term preference for quitting would view the policy as on net improving, not diminishing, the value of their take-home pay. For similar reasons, the images are distributively preferable to a tax; since smokers tend to be poorer, if there is on average a net benefit from the policy, it is actually progressive overall. It is particularly progressive compared to a tax alternative.[295]

There is, though, a fair argument in favor of subsidies or other "carrots" for smoking cessation, and likely the optimal policy is a mix of carrots and vivid images. The correlation between low willpower, smoking, and low earning potential makes smoking-cessation subsidies a highly efficient tool for redistribution. But carrots also have downsides, including the tax cost of paying for them, which is in turn compounded by the carrots' potential income effects. Both of these problems likely grow exponentially with the size of the subsidy.[296] Further, the redistributive rationale would be turned on its head to the extent that subsidies are claimed by smokers of higher income. Therefore, it would not be optimal to set subsidy levels at the full internality-correcting level (*tau*, or "τ" in Figure One)—because, as we saw before, it is better to

---

[293] AUSTRALIA GOV'T DEP'T OF HEALTH AND AGING, *supra* note 8, at 57.

[294] For surveys, *see id.*; Hammond, *supra* note 8, at 329–34. The D.C. Circuit, in finding that the U.S. FDA had failed to show evidence that vivid images would reduce smoking, cherry-picked a single sentence from Hammond's review in which he expressed reservations about the empirical methodology of one paper. *R.J. Reynolds Tobacco Co. v. Food & Drug Admin.*, 696 F.3d 1205, 1220 (D.C. Cir. 2012). Overall, he reports, "[T]he research literature unequivocally demonstrates the impact of comprehensive health warnings." Hammond, *supra* note 8, at 334.

[295] Lucas argues that "psychic" taxes are regressive, Gary Lucas, *Paternalism and Psychic Taxes: The Government's Use of Negative Emotions to Save Us from Ourselves*, 22 S. CAL. INTERDISCIPLINARY L.J. 227, 297–98 (2013), but fails to distinguish between psychic and real taxes. A tax that is collected in dollars, because of the diminishing marginal utility of money, has much greater impact on poor households. *See* Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1121 (1972).

[296] *See supra* text accompanying notes 210–211.

make multiple small mistakes than one big one. Vivid images could be used to make up the resulting gap between *tau* and the subsidy price.

Unfortunately, the optimal policy for the cue-triggered smokers in Box 1 is probably a mirror image of the ideal Box 4 strategy. Our analysis earlier suggests that taxes on inattentive smokers are probably a highly efficient revenue source, though that would be less true to the extent that they trigger labor-supply or related distortions. The revenue benefit, and its accompanying helpful income effects, have to be traded off against the sharply regressive impact of a substantial tobacco tax.[297] If the resulting optimal tax were less than *tau*,[298] the vivid images could be added to the policy mix in order to obtain full deterrence of the internality. Images are also helpful to the extent that some inattentive smokers are inattentive to the cigarette tax, but are still sensitive to vivid images (for example, because the images disrupt the tempting cues that otherwise trigger smoking).[299]

While it isn't necessarily absurd to try to enact both a tax and a subsidy at the same time, it may be more sensible simply to rely on vivid images and other "transferless" policies instead. If the tax and subsidy exactly offset—say, if all the new cigarette tax revenues are used to support smoking cessation—then on net what we've done is enact a mandate to purchase smoking cessation.[300] Smokers would also retain a marginal incentive to cut back, since they could still reduce their personal contribution to the cessation program with each puff they snuff.[301] A mandate to buy cessation could be defensible, as perhaps for some smokers the cost of cigarettes is crowding out cessation spending.[302] But it would require large administrative costs to administer both the tax and the subsidy.[303] We might justify the costs by arguing that cessation and taxing policies reached different individuals than the images alone could. If not, though, it seems better to simply rely on the images, which provide marginal incentives to quit with something like one-third the administrative apparatus.

---

[297] *See* Farhi & Gabaix, *supra* note 277, at 26 (modeling efficiency tradeoff between distortions and redistribution for low-salience taxes). While one could imagine strategies for avoiding the regressive impact of the tax—for example, implementing the "tax" as a smoking license fee, and granting free licenses to poorer households—these approaches would mostly eliminate the efficiency advantage of the tax by turning it into a *de facto* tax on income.

[298] Note that, because of the revenue benefits, the optimal Box 1 tax could conceivably exceed *tau* if there were no redistributive concerns.

[299] *See* Bernheim & Rangel, *supra* note 47, at 44–45 (suggesting that cue-triggered smokers should be unresponsive to future tax cost of smoking).

[300] *See* KAPLOW, *supra* note 178, at 13–34 (analyzing externality-correcting policy enacted together with exacting offsetting tax adjustment).

[301] *Id.*

[302] *Cf.* Susan H. Busch et al., *Burning a Hole in the Budget: Tobacco Spending and its Crowd-Out of Other Goods*, 3 APPLIED HEALTH ECON. & HEALTH POL'Y 263, 266–71 (2004) (reporting that smoking expenditures displace rent and food).

[303] For an overview of enforcement issues in tobacco excise collection, see Department of the Treasury Report to Congress on Federal Tobacco Receipts Lost Due to Illicit Trade and Recommendations for Increased Enforcement, Feb. 4, 2010, http://www.ttb.gov/pdf/tobacco-receipts.pdf.

On the other hand, if the policies don't perfectly offset, we are left with a stub version of either one, potentially with helpful economic results. For instance, if the rich pay cigarette taxes but don't take up free government cessation programs (perhaps because they can afford better programs on their own), then we have achieved a small, but quite efficient, redistributive tax.

The presence of Box 3 smokers might push us towards the three-instrument option. Although evidence is still very preliminary, some research does suggest that vivid images can prompt stronger bolstering and denial responses among the group of heavy smokers, diminishing the efficacy of the image.[304] It might be worth paying the extra administrative costs of the tax/subsidy/vivid image combination to reach these individuals, although admittedly they are a relatively small share of smokers.

If there is such a thing as a "rational smoker,"[305] their presence also would suggest that multiple instruments could be optimal. Externalities or other market failures aside, the government should not try to change the behavior of rational actors.[306] Where the government faces a mix of rational and irrational, we've seen, the *tau* or price it presents should usually be a weighted average of the zero price that should face the fully rational and whatever other prices should face various internality sufferers—essentially, balancing the harm to the rational against the help for the irrational. If government has some evidence that allows it to make educated guesses about who is in which pool, it can use multiple instruments to strike a better balance for each group.[307]

Extensive heterogeneity in smokers' need for government help, then, would offer another rationale for using multiple instruments to combat smoking. Gary Lucas argues that government should limit itself to offering opt-in commitment devices, such as voluntary smoking licenses, in order to avoid burdening possible rational smokers.[308] The trouble with this suggestion is that it leaves those who don't seek out commitment devices, our Box 3 and naïve Box 1 smokers, without any help at all. Better would be to implement a general policy for all smokers, with a relatively lower tau, that accounts for the possibility that some in that group may have no or smaller internalities. On top of that, individuals could opt into more costly policies, reflecting the fact that those who have opted in have, on average, more need for help. In other words, the voluntary license could, and likely should, exist side-by-side with vivid images.

In short, there are good reasons for the U.S. Food & Drug Administration to go forward with its stalled regulatory project on vivid images for tobacco control. And, if confronted by a skeptical court wondering why taxes and subsidies aren't preferable alternatives to infringing on the commercial speech of cigarette makers, the FDA now has some good answers.

---

[304] AUSTRALIA GOV'T DEP'T OF HEALTH AND AGING, *supra* note 8, at 41.

[305] Gary S. Becker & Kevin M. Murphy, *A Theory of Rational Addiction*, 86 J. POL. ECON. 675, 694–95 (1988).

[306] GRUBER, *supra* note 12, at 3.

[307] Galle, *supra* note 114, at 1730–34.

[308] Lucas, *supra* note 229, at 743–44. *See* Lee Anne Fennell, *Revealing Options*, 118 HARV. L. REV. 1399,1483–85 (2005), for a more detailed proposal.

VI. Conclusion

The analogy between externality regulation and internalities is powerful. The lessons of the externality literature not only help us to see why we should regulate internalities—answering, for example, the heterogeneity and information constraint objections to paternalistic regulation—but also how. But, as I have tried to explore here, internalities are also different. They present unique informational challenges we are still learning to overcome. And some standard verities of externality control, such as the clear advantages of price instruments, and the clear inefficiency of carrots, are not at all obvious when translated to internalities.

I don't mean to claim that the answers I offer here are the best or the final word. In general, my goal instead has been to sketch policy possibilities and reveal places where answers depend on unknown facts. My hope is that this work helps establish an empirical research agenda for myself and others, and to stimulate discussion about what we think we do know.

At a minimum, though, I hope that I have shown it is more than possible to make an efficiency case for nudges and other kinds of cognitively-informed regulation, particularly in the internality context. While refinements and counter-arguments certainly are likely to come, the economic case for nudging is too good to dismiss with a wave and the cry, "paternalism!"