Very preliminary draft 6/10/18

The Welfare Effects of Information

Cass R. Sunstein*

Abstract

Some information is beneficial; it makes people's lives go better. Some information is harmful; it makes people's lives go worse. Some information has no welfare effects at all; people neither gain nor lose from it. Under prevailing executive orders, agencies must investigate the welfare effects of information by reference to cost-benefit analysis. Federal agencies have (1) claimed that quantification of benefits is essentially impossible; (2) engaged in "breakeven analysis"; (3) projected various endpoints, such as health benefits or purely economic savings; and (4) relied on private willingness-to-pay for the relevant information. All of these approaches run into serious objections. With respect to (4), people may lack the information that would permit them to say how much they would pay for (more) information; they may not know the welfare effects of information; and their tastes and values may shift over time, in part as a result of information. These points suggest the need to take the willingness-to-pay criterion with many grains of salt, and to learn more about the actual effects of information, and of the behavioral changes produced by information, on people's experienced well-being.

1. Introduction

With respect to human welfare, information can be divided into three categories. Some information is beneficial; it makes people's lives go better. Some information is harmful; it makes people's lives go worse. Some information has no welfare effects at all; people neither gain nor lose from it.

Federal law often requires or authorizes federal agencies to mandate disclosure of information. Such agencies are frequently

---

* Robert Walmsley University Professor, Harvard University. In some places, this essay draws on Cass R. Sunstein, On Mandatory Labeling, With Special Reference to Genetically Modified Foods, 165 U. Pa. L. Rev. 1043 (2017). I am grateful to Cody Westfall for superb research assistance.

asked to answer four questions. (1) When should government require disclosure, or in other words, in what circumstances is there a market failure? (2) When would disclosure have desirable consequences for social welfare? (3) How can those consequences be measured? (4) When would information do more good than harm?

These questions arise in many contexts, involving calorie labels, mortgage disclosures, energy efficiency labels, fuel economy labels, credit card disclosures, labels for genetically modified food, nutrition facts panels, country of origin labels, dolphin-free tuna labels, sunscreen labels, "conflict minerals" disclosures, graphic warnings for cigarettes, and much more. Some of these labels are designed to enable consumers to protect themselves from risks, involving money or health. Some of them attempt to protect third parties or respond to moral concerns – as, for example, when labels offer information that bears on animal welfare. Some of them respond to some kind of consumer (or interest-group) demand for government action, whether or not risks are involved.

In all of these cases, assessment of welfare effects can be daunting. Sometimes government agencies know far too little to make any kind of projection of likely effects, because they do not know how people will respond. If, for example, people are informed that certain refrigerators or microwave ovens are energy-efficient, or given clear information about the costs of operating household appliances, how will they respond? In some cases, agencies simply confess their lack of knowledge. Sometimes they engage in "breakeven analysis," explaining that if the benefits reach a certain level, the costs of mandatory disclosure will be justified. Sometimes agencies are able to quantify the benefits and costs of mandatory labels, or at least significant subsets of them, either by using endpoints (economic savings, health benefits, and economic costs) or by measuring private willingness-to-pay for labels.

As we shall see, the costs of labels may be higher, and the net benefits lower, than is readily apparent, because they may produce significant decreases in consumer welfare -- as, for example, when calorie labels lead people to buy goods that are lower-calorie but less tasty, or when energy efficiency labels lead people to purchase appliances that cost less to operate but are less attractive. The point was captured in a reaction of one government official to mandatory calorie labels: "They ruined popcorn!"

As we will also see, private willingness-to-pay appears to be the best approach in theory (Viscusi, 2018); on optimistic assumptions, it should capture everything that consumers stand to gain from information. But for two principal reasons, obtaining a useful willingness-to-pay measure, capturing welfare effects, raises serious empirical, normative, conceptual challenges.

The first reason is that to be worth using, willingness to pay should be informed, and often consumers lack the information that would enable them to decide how much to pay for (more) information. The second reason is that willingness to pay, stated in advance, may fail to capture the actual welfare effects of information. In some of the relevant contexts, preferences may be labile and endogenous. Once informed about health risks associated with certain foods, for example, people might (begin to) develop different tastes. On plausible assumptions, for example, salt and sugar labels can lead to transformations in tastes. Ex ante willingness to pay figures will be insufficiently informative on that count -- which creates serious problems for welfare analysis.

There is a large issue in the background here, which involves the welfare effects of nudges more generally. Labels are a kind of nudge, understood as a choice-preserving intervention that steers people in certain directions (Thaler and Sunstein 2008). Nudges include reminders; warnings; the layout of a cafeteria or a grocery store; statements about existing social norms; and default rules. In some cases, nudges succeed in moving human behavior in the desired directions, producing significant benefits (Benartzi et al. 2017). Some of the welfare effects, both positive and negative, should not be difficult to catalogue. But suppose that people experience a welfare loss *as a result of the nudge itself*, or *as a result of the very behavior that the nudge induces*.

For example, people might not enjoy a reminder or a warning; neither of these may be pleasant to receive. Many people experience a welfare loss when they are informed about risks that they are running, even if on balance they are better off as a result of obtaining that information. The welfare effects of learning about a risk of heart disease or cancer must include consideration of the potential hedonic loss associated with that very information. In addition, the behavior change induced by nudges may include a loss of welfare; consider exercising, when exercising is not pleasant.

Of course it is also true that information, or nudges, might produce benefits rather than costs – again, not because of standard effects, but because they confer hedonic benefits. For example, people might enjoy learning for its own sake; exercising might be pleasant. But my principal emphasis here is on losses rather than gains. One of my goals is to make some progress in understanding welfare losses potentially associated with nudges, even when they are desirable on balance.

2. Costs and Benefits

With respect to mandatory disclosure of information, the initial question, of course, is whether there is some kind of market failure, justifying government intervention (Sunstein 2018). Sometimes the market is providing the optimal level of information. But for standard reasons, it may fail to do so, and unrealistic optimism or present bias on the part of consumers may aggravate the problem (Sunstein 2018).

Let us bracket these issues and assume that there is a market failure. The question remains: do the benefits of disclosure justify the costs? If it would be expensive to comply with a disclosure requirement—say, $800 million annually—the question whether the benefits are sufficient would be put in stark relief. We could easily imagine disclosure requirements that do little good, perhaps because consumers pay no attention to them. If so, such requirements would be unjustified on cost–benefit grounds. We could also imagine disclosure requirements from which consumers and third parties would benefit greatly.

2.1 Not wanting to know, and wanting not to know

It is natural to assume that receiving information is beneficial, especially for consumers; for example, it seems good to learn about product characteristics, including performance and durability. But exactly how good? There is a great deal of information that people do not care to receive (and hence has no value for them), and a great deal of information that people want *not* to receive (and hence has negative value). In some cases, people do not want to know (Ullmann-Margalit 2017); in others, they want *not* to know.

You might not much care to learn about the precise metals that were used to make your automobile, or whether the coffee beans at the local store came from Brazil, Colombia, or somewhere else. More broadly, you might not want to know whether you will get Alzheimer's disease, whether you have a genetic susceptibility to cancer and heart disease, and the year of your likely death. The general phenomenon of "information avoidance" suggests that people often prefer not to know (Golman et al. 2017). In some cases, people might be willing to pay nothing for information, and in other cases, they might be willing to pay *not* to receive information.

I conducted a small-scale study on such issues, using Amazon's Mechanical Turk, and asking about 400 Americans about whether they wanted information of these general kinds, and also how much they would be willing to pay for that information. Only 47 percent said that they want to know if they will get Alzheimer's disease. About 58 percent said that they want to know whether they have a genetic disposition to cancer or heart disease. Only 27 percent said that they want to know the year of their likely death.

With respect to information that bears on consumption choices, only 43 percent said that they want calorie labels at restaurants. Their willingness to pay for that information was modest: just $45 annually. We can reasonably infer that many people do not want to receive some information even if it seems relevant to their choices -- and that when they do want that information, they do not place a high value on it. Perhaps they think that the information would not affect their choices, or that it would be unpleasant to receive it (Golman et al. 2017).

2.2. The case of bioengineered food

To see some of the complexities, consider a proposed rule from the U.S. Department of Agriculture (2018a), the Bioengineered Food Disclosure Standard. The agency did not have much difficulty with respect to costs. It projected first-year costs of $600 million to $3.5 billion, with ongoing annual costs of between $114 million and $225 million. With respect to benefits, the agency flatly said that there would be *none* "to human health or the environment" (U.S. Department of Agriculture 2018b, p. 65). To fortify the point, it added that even if the new labels changed the ratio of bioengineered (BE) to non-BE food purchases, "there would be no impacts on human health or the environment."

Nonetheless, the agency pointed to two categories of benefits. The first involved elimination of (a) a more aggressive approach from one state (Vermont) that might drive the national market or (b) the inefficiencies of diverse state-level labeling requirements. These kinds of benefits are irrelevant to my topic here. The second and more pertinent category of benefits comes from providing consumers with "reliable information about BE food products." As the agency noted, consumers "have expressed interest in this information." But it declined to try to monetize that interest. It observed that "consumer surveys, experimental studies, and market outcomes suggest different valuations." It added that in this context, "Willingness to pay and other experimental studies for BE foods are particularly problematic as a basis for computing potential benefits (consumer surplus), with a number of researchers questioning the high negative consumer valuation of BE products resulting from these types of studies."

Those questions are perfectly legitimate. The surveys, suggesting that people are worried about BE products, may not match actual behavior with respect to purchasing decisions; there is good reason to think that they do not (Kalaitzandonakes et al. 2005). In the abstract, people might *say* that they are worried, but in stores, they show that they are not. To the extent that it exists, consumer willingness to pay for the information may well depend on a mistaken belief that BE food products are unsafe or harmful to the environment (Sunstein 2017; Sunstein 2018). If willingness to pay is based on a mistake of fact, the preferred remedy should be to correct the error rather than to require labels -- not least because labels could aggravate and spread the error, by making people think that federal regulators are concerned about the risks (Bar-Gill 2018).

In the end, the USDA (2018b, p. 8) concluded that as against a baseline of no state-level BE labels, "there are no quantified benefits associated with the Federal standard." What is noteworthy is that the USDA did not say that the benefits were zero. Instead it declined to quantify the benefits of providing the information *in any way*.

2.3 Four approaches

As the BE example suggests, U.S. agencies have often faced considerable difficulty in quantifying the costs and benefits of disclosure requirements. In fact, they have adopted four distinctive approaches, imposing increasingly severe information-gathering demands on public officials producing regulatory impact analyses. It is not always easy to explain why agencies choose one or another approach in particular cases.

The first approach—adopted in the case of BE food, and sometimes the most candid—is to confess a lack of knowledge by acknowledging that, in light of existing information, some costs and (especially) benefits simply cannot be quantified.[1] The problem with

[1] For an important decision upholding a refusal to quantify benefits, on the ground that quantification was not feasible, see Investment Co. Institute v. Commodity Futures Trading Commission (U.S. Court of Appeals for the D.C. Circuit 2013). In the

this approach is that it suggests that the decision to proceed is essentially a stab in the dark. When the stakes are large, that seems unacceptable, certainly for policymakers. It is also a disservice to the public: Should regulators impose significant costs on the private sector without making every effort to be transparent about the benefits that disclosure might confer? To be sure, quantification may turn out not to be feasible.

The second approach involves "breakeven analysis," by which agencies describe what the benefits would have to be in order to justify the costs—and suggest that the benefits are indeed likely to be of the requisite magnitude. Suppose, for example, that a mandatory label would have annual costs of $100 million and that the product is purchased, every year, by fifty million consumers. Agencies might ask: Is the label worth $2 annually to every consumer? A question of this kind might have an obvious answer.

In principle, this approach is better than a simple confession of ignorance, at least if the benefits have a lower or upper bound. Breakeven analysis is sometimes the only possible path forward. But in hard cases, it involves a high degree of guesswork, and it may be a mere conclusion, a kind of ipse dixit, masquerading as an analytic device. Without reasonable identification of lower or upper bounds, it is not so different from a confession of ignorance.

The third approach is to attempt to specify outcomes in terms of endpoints, such as economic savings or health endpoints. The advantage of this approach is that it actually points to concrete benefits, and it attempts to measure and to monetize them. But it too runs into serious difficulties. The first is epistemic: agencies may lack anything like the information that would enable them to venture such a specification. The second is that, for reasons I will explore, even an accurate specification of endpoints will not give a complete picture of the actual benefits; in crucial respects, it will almost certainly overstate

---

context of disclosure, the leading decision is National Association of Manufacturers v. SEC (U.S. Court of Appeals for the D.C. Circuit 2014), which upheld against arbitrariness review a regulation that would require disclosure of the use of "conflict minerals":

> An agency is not required "to measure the immeasurable," and need not conduct a "rigorous, quantitative economic analysis" unless the statute explicitly directs it to do so. Here, the rule's benefits would occur half-a-world away in the midst of an opaque conflict about which little reliable information exists, and concern a subject about which the Commission has no particular expertise. Even if one could estimate how many lives are saved or rapes prevented as a direct result of the final rule, doing so would be pointless because the costs of the rule—measured in dollars—would create an apples-to-bricks comparison. Despite the lack of data, the Commission *had* to promulgate a disclosure rule.

Quoting Investment Co. Institute v. Commodity Futures Trading Commission (U.S. Court of Appeals for the D.C. Circuit 2013).

them. In brief, the problem is that people might experience significant losses as well as gains as a result of receiving information (for example, if they switch to a product that is inferior along certain dimensions). An account of endpoints will ignore those losses.

The fourth approach is to identify consumers' willingness-to-pay. As a matter of abstract principle, that approach might seem to be the right one; on optimistic assumptions it should capture the full universe of losses and gains from the label (Viscusi, 2018). At the same time, it runs into serious and perhaps insuperable normative, conceptual, and empirical challenges, some of which are distinctive to the setting of willingness to pay to obtain information, some of which involve the limits of the willingness-to-pay criterion in general. The most obvious problem is that it is difficult to elicit people's *informed and unbiased* willingness-to-pay for labels. An additional challenge involves the potentially labile character of relevant preferences, including preferences for the very good for which information is provided. Note in this regard that when people offer their willingness to pay, they are attempting to solve a *prediction problem*. That problem may be difficult to solve, perhaps especially (but not only) when people are asked about whether they want to receive information.

In short, we are dealing here with a possible disjunction between "decision utility" (the utility expected at the time of decision) and "experienced utility" (the utility actually experienced) (Kahneman and Thaler 2015). For that reason, willingness to pay measures may be a very crude proxy for the actual welfare effects of obtaining information. I mean this point to raise a concern about willingness to pay for information, but it applies more broadly, for example to the valuation of morbidity risks (Viscusi, 2018). If we see willingness to pay as an effort to solve a prediction problem, we might wonder whether it is likely to be a sufficiently accurate measure of the actual welfare effects of (say) a severe concussion, chronic bronchitis, ringing in the ears, or a nonfatal heart attack. Of course it is true that more accurate measures may be unavailable.

3. Costs

On the cost side, some of the questions are relatively straightforward. Regulators may well be able to learn the total cost of (for example) producing fuel economy labels and placing them on new vehicles. In the case of BE food, the agency offered a remarkably large range for costs for the initial year, but the projections were reasonably bounded after that. A separate difficulty, which agencies have often ignored, arises *when the information itself imposes costs on consumers*. It is a mistake to ignore those costs, even if they prove difficult to quantify, and even if consumers benefit on net.[2] Those costs come in several different forms. Some of them will usually be low, but not always.

3.1 A Small Cognitive Tax

A cost is involved in reading and processing the information. For each consumer, that cost is likely to be quite low, but across a large number of purchasers, it might turn out to be significant. Information disclosure is, in a sense, akin to a paperwork burden. To be sure, consumers are not compelled to read and process what is disclosed. But even for those who seek to ignore it, its very presence may operate as a kind of cognitive tax. Because people have limited mental bandwidth, that tax may not be safely ignored. (If there is a Hell, it is filled with warnings.)

3.2 Ruining Popcorn, 1: A Hedonic Tax on Those Who Do Not Change Their Behavior

Much more importantly, the cost may be hedonic, not cognitive. Suppose that smokers are given information about the adverse health effects of smoking or that visitors to chain restaurants are given information about the caloric contents of food. Many members of both groups will suffer a hedonic loss. Consider, for example, smokers who cannot or will not quit, and customers who decide to choose high-calorie foods notwithstanding the labels. In hedonic terms, such people will lose, rather than gain, if they are miserable or at least sadder at the time of purchase.

To be sure, there is a normative question whether regulators should count, as costs, the adverse hedonic effect of truthful information. Is it a cost, or a benefit, if people learn, truthfully, that they have diabetes or cancer? On net, that might well be a benefit, at least if they can do something about the problem. (If they cannot or do not, it might be a cost, on net, at least in terms of subjective welfare.) But there is a cost as well, and potentially a large one, even if the net effect is positive. So long as we are operating within a welfarist framework, the hedonic loss must be treated as a cost. It might turn out to be low, but regulators should not ignore it (as they typically do).

Compare: Many people do not want to get blood tests, even if doctors advise them to do so, because they do not want to bear the

---

[2] For a useful discussion in an especially controversial area, see Levy et al. (2016).

hedonic cost of less-than-good results. The failure to get the tests might be a product of a behavioral bias (for example, present bias), but it might also be a product, in part, of a rational aversion to receiving negative information. Recall that large numbers of people do not want to know if they have a genetic disposition to cancer or heart disease; one reason must be the hedonic loss of receiving that information. Some labels belong in the same category, in the sense that they give people information that they are unhappy to hear (again, this is so even if on balance, they are better off with it than without it).

3.3. Ruining Popcorn, 2: A Hedonic Tax on Those Who Do Change Their Behavior

Even if people might be able to quit smoking or end up choosing lower-calorie items, and will hence benefit greatly on net, they will incur a cost by seeing something that inflicts pain. In principle, that cost should also count, even if it is greatly outweighed by benefits. The point is not that the hedonic cost is necessarily a trump card; if people make different choices once they are informed, the presumption should be that they are better off. But *by how much*?

To answer that question, the hedonic cost must be taken into account. For many people, a calorie label imposes a serious cost, simply because it informs them that the delicious cheeseburger they are about to eat is also going to make their belly bulge. It is true that there is a difference between theory and practice, and that in practice, reasonable regulators might not know what to do with this hedonic cost (other than perhaps to recognize it without quantifying it). The only point is that the cost is real.

3.4. A Consumer Welfare Loss

There is a fourth loss, in the form of a consumer welfare loss. Suppose that people decide that on balance, they should have a salad rather than a cheeseburger, on the ground that the latter has many more calories. If they choose the salad because of a calorie label, they are probably better off on balance—and in a sense, they are sadder but wiser (and healthier). They are sadder to the extent that they enjoy their meal less.

Assessment of the magnitude of the loss poses serious conceptual and empirical challenges (addressed below), but there is no question that it exists, and that it might turn out to be a significant fraction of the benefits. In principle, a decision to forego the hamburger might make people only modestly better off, if the hedonic loss is almost as high as the health gain. Whenever a mandatory label leads people to substitute product A for product B, there is a welfare loss to the extent that aside from the characteristic on which the label focusses, product B is better than product A.

Suppose, for example, that consumers are choosing between two essentially equivalent cars; that the more fuel-efficient one would cost $2000 less annually to operate because of its greater fuel efficiency; that the less fuel-efficient one would cost $500 upfront; and that because of the fuel economy label, they select the fuel-efficient car. For each such consumer, we might be tempted to say that the label has produced $1500 in gains. But in actual practice, the effects of a fuel economy label will be much more complicated to assess. Some consumers will end up purchasing cars that are more fuel-efficient but inferior along some dimension, so that they will gain $1500 minus *X*, where *X* refers to the desirable features of the unchosen car that they otherwise prefer. It is hard for public officials to know whether *X* is, on average, $100, or $1000, or $1450.

3.5 The Problem of Endogenous Preferences

All this assumes that preferences are consistent and exogenous. In some contexts, however, that assumption is not correct (Grüne Yanoff and Hansson 2009). This point complicates the foregoing analysis and creates a risk that analysis of costs will ignore shifts in tastes that are induced by labels themselves.

Suppose that at Time 1, people enjoy hamburgers a lot and enjoy salads only a little. Now suppose that having seen the labels, people switch at Time 2 because they want to make healthier choices. At Time 2, they suffer costs as a result of the switch; they miss hamburgers (delicious!) and they do not much like salad (boring!). But at Time 3, people might come to dislike hamburgers (disgusting!) and to love salad (fresh!). In principle, changes in people's preferences must be taken into account by the considered cost–benefit analysts, though doing so presents empirical serious challenges. It might be difficult to know the magnitude of the change and even the sign (perhaps those who switch to salad will crave hamburgers and grow to despise salad).

4. Benefits

On the benefits side, the assessment can be even more challenging.[3] If the government mandates a fuel economy label, for example, agencies might want to project the economic and environmental benefits from the mandate.

4.1 Behavior change

To do that, they have to identify the effect of labels on behavior. In principle, a randomized controlled trial would be valuable and perhaps necessary for that purpose. If one group sees a particular label and a similar group sees a different label (or no label), regulators should be able to specify the effect of the label on purchasing decisions. Armed with that information, they could estimate economic and environmental consequences (at least if they could generalize from the trial). Unfortunately, it is sometimes difficult or impossible to run randomized controlled trials. In these circumstances, making any kind of projection of how consumers will react to a label is exceedingly difficult. Agencies might rely on online surveys or on focus groups, which can provide relevant information.

---

[3] For example, according to the Environmental Protection Agency and the Department of Transportation (2011), speaking of new fuel economy labels,

> The agencies recognize that Executive Order 13563 directs agencies "to use the best available techniques to quantify anticipated present and future benefits as accurately as possible." In this context, however, quantitative information is not available, and the agencies have therefore chosen instead to continue with a qualitative assessment of benefits. It is difficult to develop a good baseline for the fleet using the existing label, partly because the existing label is not designed to incorporate advanced technology vehicles. It is even more difficult to develop a comparison for the fleet with the new labels, because the effects of label designs on vehicle purchases are not known. Thus, any assessment of quantitative effects of label design on vehicle sales involves a great deal of speculation. The agencies believe that informed choice is an end in itself, even if it is hard to quantify; the agencies also believe that the new labels will provide significant benefits for consumers, including economic benefits, though these benefits cannot be quantified at this time.

In short, "The primary benefits associated with this rule are associated with improved consumer decision-making resulting from improved presentation of information. At this time, EPA and NHTSA do not have data to quantify these impacts" (Environmental Protection Agency and Department of Transportation 2011).

But it would be hazardous to project, from that research, specific numbers with respect to behavior change.

4.2. Willingness to pay

An additional problem is that for the reasons given thus far, the projection would not give an adequate estimate of the (net) benefits. We have seen that if people are buying cars that are more fuel-efficient but otherwise highly undesirable, there will be a welfare loss. For that reason, regulators might explore the issue from another direction. Rather than asking about the economic savings from the fuel-efficient car, they might ask an entirely different question: *how much would consumers be willing to pay for a fuel economy label*?

In the context of reports about home energy use, Allcott and Kessler (2015) have asked a question of that kind. In their valuable and provocative work, they find that on average, people are willing to pay *something* for those reports, but that the average amount that they are willing to pay is far less than the average economic savings that people enjoy as a result of the reports. One implication is that the standard evaluation greatly overstates the net welfare gain from the reports (by a factor of five). It is not clear why the willingness-to-pay figures are so much lower than the economic gain; why would people pay (say) $2.30 for a report that would enable them to save (say) just $7?

But on reflection, the question is not so mysterious. It is plausible to speculate that the relatively lower WTP reflects an assortment of welfare losses from receiving the report: the time spent reading it, the emotional tax of receiving less than good news, the time spent taking steps to reduce energy use. Whatever we think of the precise numbers given by Allcott and Kessler, willingness to pay should capture factors of this kind. (Of course some people undoubtedly *enjoy* the report, which suggests that a full accounting would have to identify the hedonic benefits of receiving information.)

Under ideal conditions, and bracketing the endogeneity issue, the right question for regulators to ask involves willingness-to-pay. They should not focus solely on the economic benefits that consumers might receive if (for example) they purchase more fuel-efficient cars. The reason is that on optimistic assumptions, the willingness-to-pay question ought to include everything that matters to consumers. (Of course it is true that the question will not fully capture third-party effects, nor will it capture welfare effects if preferences are endogenous.)

4.3 Anchoring, information, and behavioral biases

As an empirical matter, however, it is not easy to obtain a reliable answer to the willingness-to-pay question, or anything close to it. We might simply ask people, as Allcott and Kessler did, but for their answers to be relevant, it would be important to provide pertinent information—for example, about the potential benefits (purely economic and otherwise) of labels. Unfortunately, providing that information might "anchor" consumers and hence bias their answers. Suppose, for example, that consumers were told that the average family saves $8 per year as a result of receiving home energy reports, or that fuel economy labels lead people to save $100 annually, on average, as a result of purchasing more fuel-efficient vehicles. Respondents would likely anchor on such numbers. If so, it is not clear what their answers would tell us. Anchored answers would not be especially informative about welfare effects.

Imagine that the problem of anchoring could be overcome and that informed consumers would be willing to pay (say) $10, on average, for fuel economy labels. If so, we might have some sense of the benefits, at least if behavioral biases are not distorting people's answers. In actual practice, however, such biases might well produce distortions; consider present bias and optimistic bias, which may lead to willingness-to-pay figures that are unduly low in light of the welfare benefits. In any case, survey evidence is imperfectly reliable, in part because of the familiar problems with contingent valuation studies of any kind (recall the USDA's comments about consumer surveys) and in part because of the immense difficulty of informing consumers in a sufficiently neutral way.

For health-related disclosures, the problem is even harder. One goal of calorie labels, for example, is to reduce obesity, which causes an assortment of health problems, including premature mortality. Regulators have established ways to turn health-endpoints into monetary equivalents. For example, a statistical death is now valued at about nine million dollars.[4] But how many premature deaths would

[4] The defining work here comes from W. Kip Viscusi. See Viscusi (2018); many people draw on his research. See, e.g., Thomson and Monje (2015) explaining, "On the basis of the best available evidence, this guidance identifies $9.4 million as the value of a statistical life." See also Sunstein (2014) providing the underlying theory and a discussion of how "agencies . . . assign monetary values to the human lives that would be saved by a proposed regulation."

be prevented by calorie labels? And what would be the effect of such labels on adverse health outcomes short of death?

To answer such questions, regulators have to undertake two tasks. First, they must begin by making some prediction about the effect of calorie labels on what people choose to eat. Second, they have to follow that prediction by specifying the health consequences of lower levels of caloric intake. At least it can be said that if they can accomplish those tasks, they will have some sense of the benefits of the labels, once (and this is a third task) they turn the various consequences into monetary equivalents. After undertaking all three tasks, regulators will have specified endpoints—but for the reasons given, a specification of endpoints will overstate benefits because it will not include various cognitive and hedonic losses. (As noted, there be also be hedonic benefits.)

Alternatively, we could (again) ask how much people would be willing to pay for calorie labels.[5] As before, asking that question is, in principle, preferable to an effort to assess health-states, because the answer will capture all variables that matter to consumers.[6] Also, as before, there are formidable challenges in using surveys to elicit reliable numbers free from biases of various kinds. And if preferences are endogenous and labile, willingness to pay numbers might greatly understate the welfare gain from labels. Recall that people might develop tastes for the products to which they shift. (I am also bracketing the questions raised by addictive goods, such as cigarettes, for which labels might be beneficial on welfare grounds precisely because they help break the hold of the addiction. Note that

---

[5] See Loureiro et al. (2006, p. 263) finding that "on average, consumers are willing to pay close to 11 per cent above the initial price to obtain cookies with nutritional labelling." Further, "Consistent with prior expectations, our results also indicate a difference between the [willingness-to-pay] of individuals suffering from diet-related health problems (estimated mean 13 per cent) and those who do not suffer any diet-related health problems (estimated mean 9 per cent)" (Loureiro et al. 2006, p. 249).

[6] In the words of the FDA (2014, p. 64),

> To our knowledge, Abaluck (2011) is the only study that translates the potential effect of increasing nutrition information on consumption into estimates of welfare gains using willingness-to-pay based on revealed preferences (Ref. 43). This study uses the variation in nutrition information generated by Nutrition Labeling and Education Act (NLEA) as a method to determine how changes in individuals' beliefs about nutrient content affect consumption decisions. The differential changes in nutrition information across food categories, measured in units of calories per gram, allow the study to identify a general model of food demand as a function of nutrient characteristics that accounts for the total daily diet, prior beliefs about nutrient content, and preferences, including willingness to substitute across food categories.

As before, however, the willingness-to-pay criterion may run into normative objections, even from the standpoint of welfare. See generally Bronsteen et al. (2015) raising questions about willingness-to-pay in view of people's occasional failure to know what will promote their welfare.

cigarette taxes appear to make smokers happier (Gruber and Mullainathan 2002).)

4.4 Willingness to pay as a solution to a prediction problem

There is a larger problem, to which I have briefly referred. When consumers state their willingness to pay, they are solving a *prediction problem*. To take a mundane case: When a consumer decides to spend $30,000 for a new car, rather than $25,000 or $35,000, she must be making a prediction about the welfare effects of the expenditure. In choosing among three cars, the solution to a prediction problem, for individual consumers, is not exactly easy. (Can consumers reliably foresee the welfare effects, for them, of each of the cars?) Turn now to health risks: a broken back, a severe concussion, heart disease, diabetes. Deciding how much to pay to eliminate a 1/n risk of any one of those requires a judgment about what it would be like to suffer from them. Are consumers in a good position to make that judgment? Usually not. Are they in a good position to make judgments about much it is worthwhile to obtain information that would permit them to reduce the risk of suffering from those conditions? Usually not.

In light of these challenges, regulators have two highly imperfect options. First, they can work on the two relevant tracks to try to produce answers: exploring end-points and enlisting surveys. On prominent occasions, they have tried the former.[7] Second, they can acknowledge the difficulties, confess that they cannot surmount them, and use "breakeven analysis," by which they ask what the benefits would have to be in order to justify the costs, and then do what they can to generate a reasonable lower bound.

---

[7] See U.S. Food and Drug Administration (2011) noting the longer lifespans, fewer cancers and diseases, as well as increased property and monetary values of non-smokers. See also U.S. Department of Labor (2016) requiring that employees have access to OSHA logs, and U.S. Environmental Protection Agency and U.S. Department of Transportation (2011) explaining, "The agencies believe that informed choice is an end in itself, even if it is hard to quantify; the agencies also believe that the new labels will provide significant benefits for consumers, including economic benefits, though these benefits cannot be quantified at this time." Finally, see U.S. Food and Drug Administration (2014, p. 11) explaining, "The final rule may also assist consumers by making the long-term health consequences of consumer food choices more salient and by providing contextual cues of food consumption."

Suppose, for example, that an energy-efficiency label for refrigerators would cost $10 million annually and that eight million refrigerators are sold in the United States every year. Even if the average consumer saves only $0.50 annually as a result of the label, the cost will be made up in just three years. Breakeven analysis can be crude, but in some cases, it will suggest that the argument for labels is either very strong or very weak.

5. Third Parties—and Morality

Some actual or imaginable labels are meant to protect third parties, not consumers as such. Suppose that some or many consumers are concerned about the use of certain minerals to finance mass atrocities, and they favor labeling, or some kind of disclosure requirement, so that consumers can decline to purchase products that contain such minerals. Or suppose that consumers care about where goods were made, perhaps because they want to purchase products from their own nation or perhaps because they do not want to purchase products from nations that do not respect human rights. They might seek "country of origin" labels for that reason. Or suppose that some or many consumers care about the welfare of animals in general or certain animals in particular; because they do, they seek labels to reflect how animals were (mis)treated.

In such cases, there are two sets of benefits: (1) the benefits to consumers themselves, assuming that they would enjoy a welfare gain if their moral commitments were vindicated (Posner and Sunstein 2017) and (2) the benefits to third parties. The two are separate. In principle, the right measure of (1) should be willingness to pay (ibid.), but it will not be simple to elicit it.

In some of these cases, the third-party effects are not obscure, and the real challenge is how to quantify them. As before, it is necessary to begin by making some projections about consumer behavior. To what extent would consumers change their purchasing habits in response? Even if that question can be answered, it would be necessary to tie any such changes to reduced costs or increased benefits for third parties. And even if that problem can be resolved, it would be necessary to quantify and monetize the resulting effects. It is no wonder that in the context of conflict minerals, the agency concluded that quantification was not possible.[8] Perhaps it should have engaged in some form of breakeven analysis, explaining that the requirement was likely to survive cost–benefit analysis even if its

---

[8] See U.S. Court of Appeals for the D.C. Circuit (2015) explaining, "The Commission was 'unable to readily quantify' the 'compelling social benefits' the rule was supposed to achieve: reducing violence and promoting peace and stability in the Congo," quoting Securities and Exchange Commission (2012).

effect were modest. But perhaps it lacked the information that would have allowed it to make that analysis plausible.

6. Taking Stock

In numerous contexts, Congress has required or authorized federal agencies to impose disclosure requirements. In all those contexts, executive agencies are required, by executive order, to catalogue the benefits and costs of disclosure requirements, and to demonstrate that the benefits justify the costs. Such agencies face persistent challenges in projecting benefits, and they use four different approaches: a refusal to do so on the ground that quantification is not feasible; breakeven analysis; projection of end-states, such as economic savings or health outcomes; and estimates of willingness-to-pay for the relevant information.

Each of these approaches raises serious questions and runs into strong objections. On optimistic assumptions, the right question generally involves willingness-to-pay (Viscusi, 2018). But in practice, people often lack enough information to give a sensible answer to the question how much they would be willing to pay for (more) information. (How much would you be willing to pay for information about the presence of chemical XYZ in your favorite food, when you know little or nothing about chemical XYZ or its effects?)

We have also seen that when preferences are labile or endogenous, even a sensible answer to the willingness to pay question may fail to capture the welfare consequences, because people may develop new tastes and values. Willingness-to-pay figures are an effort to solve a prediction problem, and in some cases, people are not in a good position to produce sensible solutions. In these circumstances, breakeven analysis is the very least that should be required, and it is sometimes the most that agencies can do. If it is accompanied by lower or upper bounds, a breakeven analysis will sometimes show that mandatory disclosure is justified on welfare grounds—and sometimes that it is not.

The challenge is that breakeven analysis is a confession of ignorance, and that without lower or upper bounds, that form of analysis will leave us at sea. It would be far better for agencies to make progress in answering difficult questions about the actual effects of information on people's experienced well-being. The next generation of work on disclosure requirements – and regulatory benefits in general – should make it a priority to produce those answers.

References

Allcott, H., Kessler, J. B. (2015). The welfare effects of nudges: A case study of energy use social comparisons. NBER Working Paper 21671. National Bureau of Economic Research. http://www.nber.org/papers/w21671.

Bar-Gill, O., Schkade, D., Sunstein, C. R. (2018). Drawing false inferences from mandated disclosures. *Behavioural Public Policy*, doi:10.1017/bpp.2017.12.

Benartzi, S., Beshears, J., Milkman, K. L. (2017). Should governments invest more in nudging? *Psychological Science*, 28(8), 1041.

Bronsteen, J., Masur, J. S., Buccafusco, C. (2015). *Happiness and the law*. Chicago: University of Chicago Press.

Golman, R., Hagmann, D., Loewenstein, G. (2017). Information avoidance. *Journal of Economic Literature*, 55(1), 96.

Gruber, J., & Mullainathan, S. (2002). *Do cigarette taxes make smokers happier?* NBER Working Paper 8872. National Bureau of Economic Research. http://www.nber.org/papers/w8872.

Grüne-Yanoff, T., & Hansson, S. O. (Eds.). (2009). Preference change: Approaches from philosophy, economics and psychology 4. New York: Springer.

Kahneman, D., & Thaler, R. (2006). Anomalies: Utility maximization and experienced utility. *The Journal of Economic Perspective*, 20(1), 221.

Kalaitzandonakes, N., Marks, L., Vickner, S. S. (2005). Sentiments and acts towards genetically modified foods. *International Journal of Biotechnology,* 7(1). https://www.inderscienceonline.com/doi/pdf/10.1504/IJBT.2005.006452.

Levy, H. G., Norton, E. C., Smith, J. A. (2016). Tobacco regulation and cost-benefit analysis: How should we value foregone consumer surplus? NBER Working Paper 22471. National

Bureau of Economic Research. http://www.nber.org/papers/w22471.pdf.

Loureiro, M. L., Gracia, A., Nayga, R. M. (2006). Do consumers value nutritional labels? *European Review of Agricultural Economics*, 33(2), 249, 263.

Posner, E. A., & Sunstein, C. R. (2017). Moral commitments and cost-benefit analysis. *Virginia Law Review*, 103, 1809.

Sunstein, C. R. (2014). *Valuing life: Humanizing the regulatory state.* Chicago: University of Chicago Press.

Sunstein, C. R. (2017). Mandatory labeling, with special reference to genetically modified foods. *University of Pennsylvania Law Review*, 165, 1043.

Sunstein, C. R. (2018). *The cost-benefit revolution*. Cambridge: MIT Press.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven: Yale University Press.

Thomson, K., Monje, C. (2015). Guidance on treatment of the economic value of a statistical life in U.S. Department of Transportation analyses. Memorandum. U.S. Department of Transportation. https://perma.cc/C6RQ-4ZXR.

Ullmann-Margalit, E. (2017). *Normal rationality*. Oxford: Oxford University Press.

U.S. Department of Agriculture. (2018a). National bioengineered food disclosure standard. Proposed rule. https://www.regulations.gov/document?D=AMS-TM-17-0050-0004.

U.S. Department of Agriculture. (2018b). Proposed national bioengineered food disclosure standard. Regulatory impact analysis. Report AMS-TM-17-0050-0015, 8, 65. https://www.regulations.gov/document?D=AMS-TM-17-0050-0015.

U.S. Department of Labor. (2016). Improve tracking of workplace injuries and illnesses. *Federal Register*, 81, 29628.

U.S. Environmental Protection Agency, & U.S. Department of Transportation. (2011). Revisions and additions to motor vehicle fuel economy label. *Federal Register*, 76, 39517.

U.S. Food and Drug Administration. (2011). Required warnings for cigarette packages and advertisements. *Federal Register*, 76, 36719.

U.S. Food and Drug Administration. (2014). Food labeling: Nutrition labeling of standard menu items in restaurants and similar retail food establishments. Regulatory impact analysis. Report FDA-2011-F-0172, 11, 64. https://www.fda.gov/downloads/Food/IngredientsPackagingLabeling/LabelingNutrition/UCM423985.pdf.

U.S. Court of Appeals for the District of Columbia Circuit. (2013). *F.2d*, 720, 372–375.

U.S. Court of Appeals for the District of Columbia Circuit. (2014). *F.3d*, 748, 369.

U.S. Court of Appeals for the District of Columbia Circuit. (2015). *F.3d*, 800, 547.

U.S. Securities and Exchange Commission. (2012). Conflict minerals. *Federal Register*, 83, 56350.

Viscusi, W.K. (2018). *Pricing Lives*. Princeton: Princeton University Press.