DOI: 10.1111/joes.12453

ARTICLE

JOURNAL OF ECONOMIC SURVEYS WILEY

A review of nudges: Definitions, justifications, effectiveness

Luca Congiu[1,2] | **Ivan Moscati**[3,4]

[1] Department of Economics, University of Insubria, Varese, Lombardy, Italy

[2] Department of Economics, Bocconi University, Milan, Lombardy, Italy

[3] Department of Economics, University of Insubria, Varese, Lombardy, Italy

[4] Centre for Philosophy of Natural and Social Science, London School of Economics, London, UK

Correspondence
Ivan Moscati, Department of Economics, University of Insubria, Varese, Italy.
Email: ivan.moscati@uninsubria.it

Abstract

In 2008, the behavioral economist Richard Thaler and the legal scholar Cass Sunstein published a book in which they advocated a novel approach to public policy based on the notion of a "nudge." Roughly speaking, a nudge is an intervention in the decisional context that steers people's decisions by acting on their cognitive biases. The notion of a nudge generated an intense debate across different disciplines and proved popular with many policy makers around the world. The present article reviews the debate and research on nudges by focusing on three main dimensions: (1) the exact definition of nudges; (2) the justification of nudge policies, with a focus on "libertarian paternalism"; and (3) the effectiveness of nudges, both over time and in comparison with standard policies.

KEYWORDS
behavioral welfare economics, boosts, bounded rationality, libertarian paternalism, nudge

JEL CLASSIFICATION
D01, D04, D90, G40, M31

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Journal of Economic Surveys* published by John Wiley & Sons Ltd

1 | INTRODUCTION

1.1 | Setting the stage

The models of individual decision making that dominated economics until the late 20th century, and that are often labeled as "neoclassical models," assume that decision-makers are "rational." Roughly speaking, by rational it is meant that decision-makers have preferences that satisfy certain requirements of mutual consistency (i.e., requirements of the form "if option A is preferred to option B, then option C should be preferred to option D"); that preferences are stable both over time and across decision contexts; that decision-makers are capable of processing information in an appropriate way; and that they choose the best option according to their preferences and information.

The descriptive validity of the neoclassical models and the normative soundness of the notion of rationality attached to them had been called into question at least since the 1950s (see, e.g., Allais, 1953; Ellsberg, 1961; Simon, 1955; Strotz, 1955). However, it was not until the late 1970s, after the accumulation of robust experimental evidence documenting systematic violations of received decision theories, that a series of new models aimed at accounting for the observed violations began to appear in economics journals (see, e.g., Handa, 1977; Kahneman & Tversky, 1979; Loomes & Sugden, 1982; Thaler, 1980). These models relied on psychological assumptions more plausible than the assumptions incorporated in the neoclassical notion of rationality, and in the 1990s they began to be called "behavioral models."

In most behavioral models, decision-makers still attempt to satisfy their preferences, but the mutual-consistency requirements that these preferences should satisfy are looser than the neoclassical ones. Moreover, it is admitted that preferences often change over time and across contexts, and that decision-makers process information in ways that are different from those supposed by neoclassical models. Following the usual terminology, we shall refer to these deviations from neoclassical rationality as "biases" or "rationality failures," and call agents not conforming to the neoclassical notion of rationality as "boundedly rational" agents.

In the early 2000s, behavioral economists began to explore the consequences for welfare economics and public policy of the abandonment of the neoclassical notion of rationality. Simply put, in the neoclassical framework economic policies producing outcomes that are ranked as highly as possible according to the decision-makers' rational preferences seemed justified. However, if the decision-makers' preferences are internally inconsistent, unstable over time, and context-dependent, and if decision-makers process information in idiosyncratic ways, the traditional rationale for economic policies collapses. Discussions about how to design and justify economic policies when the assumption of neoclassical rationality is discarded—an issue sometimes called the "reconciliation problem" (McQuillin & Sugden, 2012)—have created a new field called "behavioral welfare economics."[1]

Behavioral welfare economics received a significant boost in 2008, when the behavioral economist Richard Thaler and the legal scholar Cass Sunstein published a book titled *Nudge: Improving Decisions About Health, Wealth and Happiness*. The book was a popular success, generated an intense academic debate well beyond economics, and proved inspirational for many policy makers around the world who began to implement nudge-based policies. In 2009, Thaler and Sunstein published a revised and expanded edition of the book, and here we shall refer to it.

By "nudge" Thaler and Sunstein mean any intervention in the "choice architecture," that is, in the structure of the decision context, that "alters people's behavior in a predictable way without

forbidding any options or significantly changing their economic incentives," with the intervention remaining "easy and cheap to avoid" (Thaler & Sunstein, 2009, p. 6). Moreover, Thaler and Sunstein add, nudges "alter the behavior of Humans [i.e., individuals not complying with the requirements of neoclassical rationality] even though they would be ignored by Econs [i.e., neoclassically rational agents]" (p. 8). Nudges can be implemented either by governments or private companies. The archetypal example of a nudge is the positioning of healthy foods in cafeteria lines in a way that increases their prominence and accessibility, thereby making it more likely that the customer selects them. This is a limited change in one's decisional context that does not exclude access to other and possibly less healthy options, and that exploits the fact that people's choices appear to be greatly influenced by the order and position of the items in the choice set. Such an influence is implicitly ruled out by neoclassical requirements about the mutual consistency of preferences.

1.2 | This article

As mentioned, Thaler and Sunstein's book and their nudge approach to public policy have generated an intense debate. In the present article we review this debate by focusing on three main dimensions: (1) the exact definition of nudges (Sections 2–4); (2) the justification of nudge policies (Sections 5–8); and (3) the effectiveness of nudges (Section 9).

More precisely, Section 2 critically reconsiders Thaler and Sunstein's original definition of a nudge, while Section 3 discusses some alternative definitions of the concept. Section 4 distinguishes three types of nudges, namely, pro-self nudges, pro-social nudges, and commercial nudges that benefit both the retailer and the customer. Section 5 presents Thaler and Sunstein's justification of pro-self nudges, which is based on what they call "libertarian paternalism," while Section 6 reviews the criticisms of libertarian paternalism and the alleged justification for pro-self nudges it would provide. Section 7 discusses some justifications for pro-self and pro-social nudges that are not based on libertarian paternalism. Section 8 briefly presents the concept of "boosts," which are a type of behavioral policy that we consider complementary to nudges. Section 9 surveys the empirical evidence about the effectiveness of nudge policies, that is, the question of whether they work, under what conditions, and for how long, also in comparison to traditional policies. Section 10 concludes.

Before closing this introduction, a quick comparison between our review of nudges and other existing surveys is in order. Marchiori et al. (2017) examine the nudge literature from a psychological perspective, and do not address the definition of a nudge and libertarian paternalism. Lin et al. (2017) consider the effectiveness and ethical use of nudges in the health and medical sector but, again, leave aside debates on the nudge definition and libertarian paternalism. The essay closest to ours is Barton and Grüne-Yanoff's introduction to a special issue on nudges of the *Review of Philosophy and Psychology* (2015). However, Barton and Grüne-Yanoff appeared too early to address a series of important questions in the post-2015 literature that are related to the definition and effectiveness of nudge policies and focus on the libertarian-paternalistic justification of nudges while devoting less room to the discussion of alternative justifications.

Thus, we submit, our review not only accounts for the more recent literature on nudges, but it reconstructs it by considering all the main dimensions of the discussion (definition, justification, effectiveness). In addition, our review builds on numerous examples of nudge applications in diverse policy sectors, and hence may be of interest to both academics and practitioners.

2 | THE ORIGINAL DEFINITION OF A NUDGE AND ITS PROBLEMS

One of the earliest issues raised by nudge scholars concerns the very notion of a nudge. Given our purposes, it may be useful to repeat the definition of a nudge offered by Thaler and Sunstein (2009, pp. 6 and 8) and separate its four components. For Thaler and Sunstein, a nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without [1] forbidding any options or [2] significantly changing their economic incentives. To count as a mere nudge, the intervention must be [3] easy and cheap to avoid"; moreover, a nudge "[4] significantly alters the behavior of Humans even though it would be ignored by Econs."

Each component of Thaler and Sunstein's definition has been discussed and criticized.

2.1 | Adding options

First, consider "without forbidding any options." This clause states that a nudge cannot reduce the decision-maker's choice set (Saghai, 2013). However, it does not clarify whether a policy intervention can expand the choice set and still be classified as a nudge. In effect, experimental evidence shows that a decision-maker can be influenced by the addition of options in the choice set, specifically, by the addition of dominated options or "decoys" (Hansen, 2016).

A famous example is discussed by Kivetz et al. (2004) and Ariely (2009) with respect to different annual subscriptions to a journal. Subjects were confronted with the choice set {online subscription for \$59, print subscription for \$125, online and print subscription for \$125}. The second alternative is dominated by the third alternative (in the sense that the second costs the same as the third but offers less) and is thus termed a "decoy." A basic consistency condition of neoclassical choice theory, sometimes called Chernoff's axiom (Chernoff, 1954) or Sen's property α (Sen, 1970), prescribes that including or eliminating the decoy from the choice set should not change the individual's choice between the first and the third option. In effect, what Kivetz et al. (2004) and Ariely (2009) found is that the final distribution of the subjects' selections changed significantly depending on whether the decoy was present or not. More precisely, the share of experimental subjects that chose the third, most expensive option was higher when the decoy was present and lower when it was removed.

Therefore, experimental evidence shows that the addition of a dominated option is an aspect of the choice architecture that can steer decisions in predictable directions without "forbidding options," and that it "would be ignored by Econs" while affecting Humans. Hence, the decoy can be considered as a nudge.

2.2 | Beyond economic incentives

Regarding point two of the definition, according to which nudges should not significantly change "people's economic incentives," some scholars have suggested that this wording does not consider other relevant forms of incentives, such as time, social sanctions, or physical threats (Hansen, 2016; Hausman & Welch, 2010). For example, threatening to administer an electric shock to deter or incentivize certain conduct does not change people's *economic* incentives, but regarding it as a nudge would be implausible.

To be fair, Thaler and Sunstein acknowledge that some of their nudges "impose cognitive (rather than material) costs" and specify that "nudges count as such … only if *any costs* are low" (Thaler & Sunstein, 2009, p. 8, footnote, emphasis added). Given that the cost of an electric shock can be regarded as very high, this and similar forms of intervention should be automatically excluded from the nudge "toolkit." However, this additional criterion was not directly included in the original definition of nudge, thus creating room for ambiguity.

2.3 | Opaque and hard-to-avoid nudges

Part three of the definition refers to the easy avoidability of nudging ("easy and cheap to avoid"). The idea is that a nudge is not a mandate and one is not forced to comply with it. Thus, one can in some sense choose whether to "accept" the nudging or not (Sunstein, 2015). This argument, however, implicitly assumes that the decision-maker has preferences that are strong enough not to be influenced by the nudge and/or that he engages in a deliberation process to choose the course of action to pursue (Hansen & Jespersen, 2013; Saghai, 2013). For instance, the decoy effect disappears, or is severely reduced, when the decision-maker has a strong preference for either non-dominated option compared to when the decision-maker is indifferent between the two (Sürücü et al., 2019). Moreover, the decoy effect disappears when the decision-maker is given the proper time and incentives to revise his choice, implying that the effect of this nudge may not survive cognitive deliberation (Gaudeul & Crosetto, 2019).

However, it would be unrealistic to assume that any decision-maker always has strong preferences and engages in deliberation in any situation. For example, it is difficult to imagine that the typical citizen has clear preferences on the proper allocation rate for his pension fund. Plausibly, he lacks the cognitive ability or knowledge needed to perform the computations required to identify the preferred option. That may explain why, when proposed with a certain default allocation, the decision-maker tends to adopt it (Cronqvist et al., 2018; Thaler & Benartzi, 2004).

The possibility that the decision-maker engages in cognitive deliberation when confronted with a nudge is made even more remote by the fact that several nudges do not target the reflective cognitive system, thereby circumventing deliberation entirely (Hansen & Jespersen, 2013; Saghai, 2013; Sunstein, 2016). Supporters of "dual-system theories," such as Kahneman (2011), propose that the mind be thought of as comprising an automatic, quick, and effortless cognitive system (named System 1), where heuristic-based thinking operates, and a reflective, slow, and effortful cognitive system (named System 2), where cognitive deliberation takes place. Based on this distinction, Hansen and Jespersen (2013) and Sunstein (2016) distinguish between "Type 1" nudges, which target System 1 and hence bypass cognitive deliberation, and "Type 2" nudges, which target System 2 and hence trigger cognitive deliberation. For example, visual illusions for traffic control, such as fake potholes, are Type 1 nudges because they entail an automatic response (e.g., slowing down) to a visual cue (e.g., the pothole), and this without triggering any deliberation process.

Further, the avoidability of a nudge may depend also on its degree of transparency (Bovens, 2009; Hansen & Jespersen, 2013). "Transparency" entails that a decision-maker is aware of a nudge's existence, workings, and intention. For example, a "look right, look left" sign in a London street is a transparent nudge because any decision-maker is expected to understand that the sign entails an influence attempt (existence) that works by suggesting where to look before crossing the street (workings) for the purpose of reducing the likelihood of being hit by a vehicle (intention). But not all nudges are transparent. For instance, consider again the arrangement of food in

a cafeteria line: How many customers are expected to know that the food has been arranged to give more prominence to healthy products (existence), that it works because people typically fill their plate with what is encountered first in the line or is more salient (workings), and that the aim of the food arrangement is to increase healthy food intake among diners (intention)?

In conclusion, the conditions that would make a nudge "easy and cheap to avoid," namely its transparency and capacity to trigger deliberation, are often violated.

2.4 | Rational or boundedly-rational channels?

Finally, consider point four of the definition of a nudge, which deals specifically with the distinction between nudges and standard interventions. As discussed, an intervention counts as a nudge if it affects Humans while being ignored by Econs. For example, the food arrangement in a cafeteria line is a nudge because the order or salience of alternatives would not affect an Econ's choices, while it affects those of a Human. In contrast, a ban or fine to reduce smoking is a standard policy because an Econ would factor in the increased cost of smoking in his decision process.

However, this criterion may not be met in cases where both rationality and bounded rationality are targeted simultaneously. As an example, consider a five-cent tax on disposable shopping bags (Homonoff, 2018). This tax attaches a small monetary cost to disposable bags to decrease their demand and, as a result, pollution from plastic. On the one hand, this tax modifies the monetary cost of disposable bags. Since it is implausible that a monetary cost "would be ignored by Econs," this intervention should qualify as a standard policy, rather than a nudge. On the other hand, the effectiveness of this policy can also be explained by incorporating cognitive biases into the picture, namely, loss aversion (Kahneman & Tversky, 1979) and a salience effect (Bordalo et al., 2013).

In fact, as Homonoff (2018) shows, while the five-cent tax almost halves the demand for disposable bags, reframing this tax as an incentive, for instance as a five-cent bonus for adopting reusable bags, does not lead to any significant effect. This can be explained by loss aversion: with respect to the reference point where the bags are free, the tax is considered a loss, while the bonus is considered a foregone gain, and losses loom larger than gains. In addition, the effectiveness of this tax may also be due to its "salience" (Bordalo et al., 2013; Chetty et al., 2009), since it is applied to a good that is not the true object of the purchasing activity but is rather instrumental to it and demanded separately. Besides, this salience effect might be amplified by the fact that it involves a good that was previously free (Shampanier et al., 2007).

In sum, the five-cent tax on disposable bags reduces their demand by acting simultaneously on a rational channel (cost modification) and on boundedly-rational channels (loss aversion, salience effect). Hence, according to the definition of a nudge, this intervention cannot be classified as a nudge because it would not be ignored by an Econ. Still, it would be difficult to classify this intervention as a standard policy, as it also—and primarily—targets cognitive biases, and its effectiveness seems to depend crucially on this circumstance.

3 | REVISED DEFINITIONS OF A NUDGE

The original definition of a nudge has gone through a series of refinements or modifications aimed at improving it (e.g., Congiu & Moscati, 2020; Hansen, 2016; Hausman & Welch, 2010; Mongin & Cozic, 2018; Saghai, 2013).

3.1 | Hansen's definition

Arguably, the most articulated discussion of the notion of a nudge is that provided by Hansen (2016). At the end of his analysis, Hansen defines a nudge as an intervention that exploits rationality failures in an instrumental way:

> A nudge is … any attempt at influencing people's judgment, choice or behaviour in a predictable way which works by making use of [people's] boundaries, biases, routines and habits as integral parts of such attempts (p. 4; a similar definition can be found in Mongin & Cozic, 2018, p. 2).

In particular, for Hansen, (1) the instrumental use of rationality failures may include the addition of options to the choice set; this inclusion supersedes problem 1 of the original definition; (2) insofar as an intervention exploits at least in part rationality failures, it is a nudge even if it relies on "changing incentives [and on] the provision of factual information or rational argumentation" (Hansen, 2016, p. 4). Criterion (2) turns the five-cent tax into a nudge because, even though it relies on rational drivers (cost modification), it also relies on cognitive biases and boundaries (loss aversion and salience).

Still, as Mongin and Cozic (2018) have noted, criterion (2) remains problematic because it is not always obvious whether an intervention truly exploits cognitive biases or whether its effectiveness can be explained on purely rational grounds. We have already discussed the difficulty of separating rational and boundedly-rational channels in the case of an intervention changing incentives, specifically the case of a five-cent tax on disposable bags (Section 2.4). Let us now discuss this issue in the case of information provision. Consider the following message: "Most of your peers have successfully quit smoking." This message may be interpreted as targeting only rationality, in that it signals a lower cost or higher success rate of quitting smoking than one anticipates. In this case, according to Hansen's criterion (2), the message would qualify as a standard policy rather than a nudge. On the other hand, the message might also work because it sets a norm that creates social pressure for quitting smoking ("If I quit smoking, I can fit in with my group of peers"). Insofar as one considers a preference for social conformity as hardly compatible with neoclassical rationality, for it easily makes the individual's preferences over options mutually inconsistent and/or unstable, the message would indeed qualify as a nudge (for a review of psychological research on social conformity, see Cialdini & Goldstein, 2004).

At any rate, the point is that it is far from obvious to state whether the message at issue, and more generally any other policy intervention, targets only rational drivers or whether it works because it exploits some rationality failures. Consequently, the classification of an intervention as a nudge or a standard policy remains problematic, even though in this respect Hansen's definition is less problematic than Thaler and Sunstein's original definition.

3.2 | Sunstein's definition

While Hansen's (2016) modified definition of nudges connects them to rationality failures, other authors have adopted a different strategy to address the flaws in the original definition. This strategy focuses on the idea that nudges should preserve the decision-maker's liberty and autonomy, and is epitomized by Sunstein's updated definition of nudges as:

> Private or public initiatives that steer people in particular directions but that also *allow them to go their own way* (Sunstein, 2018, p. 61, emphasis added).

Berthet and Ouvrard (2019) refer to this definition as a nudge in the "broad sense," opposed to the "narrow sense" of the definition by Hansen (2016), discussed above. In fact, this definition does not refer to the specific workings of a nudge, such as the relationship with cognitive biases, thereby also including in the family of nudges interventions that appear wholly compatible with neoclassical rationality. A famous example in the nudge literature is that of GPS, which suggests a particular route without forcing the driver to follow it (Sunstein, 2015, 2018). Similarly, educational campaigns may convey sheer factual information and so not infringe on an individual's liberty and autonomy, thus qualifying as nudges in Sunstein's broad sense.

When considered from the viewpoint of libertarian paternalism, which prescribes that an intervention should preserve freedom of choice (see Section 5), Sunstein's broad definition allows the separation of libertarian-paternalistic interventions from hard-paternalistic ones, such as bans and mandates. However, since Sunstein's definition does not refer to the specific channels through which nudges operate, that is, rational versus boundedly-rational channels, it is often of little use for distinguishing nudges from standard policies.

4 | TYPES OF NUDGES

There is a final important issue related to the definition of a nudge that should be addressed. The various definitions discussed so far highlight either the instrumental use of rationality failures (Hansen's narrow definition) or the liberty-preservation aspect of the intervention (Sunstein's broad definition). Yet, nowhere in these definitions is it specified who benefits, or should benefit, from the nudge. When this aspect is taken into account, it is possible to distinguish three types of nudges: nudges that benefit the nudged individual; nudges that benefit the society; and nudges that primarily, but not exclusively, benefit the nudgers.

4.1 | Pro-self nudges

The first type of nudges is aimed at benefitting the nudged person. Hagman et al. (2015, p. 441) call these nudges "pro-self," and define them as nudges that "help individuals steer away from irrational behavior … which decreases their long-term well-being." Examples of pro-self nudges are the food arrangement in the cafeteria, aimed at promoting the customer's health, and a default allocation to a retirement fund, aimed at increasing the worker's savings.

As will be discussed in Section 5, libertarian paternalism authorizes policies that increase the targeted person's well-being, and therefore it validates pro-self nudges.

4.2 | Pro-social nudges

The second type of nudges is primarily aimed at increasing the welfare of society by steering the nudged persons away from behaviors that would reduce the common good. These nudges are called "pro-social," or "pro-others," or simply "social" nudges (Hagman et al., 2015; Hands, 2020, 2021; Nagatsu, 2015). Examples are nudges to foster tax compliance and energy

conservation through social comparison, such as when tax debtors are told that their peers pay their taxes on time (Behavioural Insights Team, 2016; John & Blume, 2018; Kettle et al., 2016) or households receive feedback on the lower energy consumption of their eco-friendly neighbors (Allcott, 2011; Allcott & Rogers, 2014; Bernedo et al., 2014). Such interventions hardly benefit the decision-maker primarily or directly: while more taxes imply a more efficient society, and less energy waste means a cleaner planet to live on, it would be difficult to argue that the decision-maker benefits the most from these interventions. Similarly, nudges aimed at raising donations for charitable causes by using defaults and decoys (Pittarello et al., 2019; Rubaltelli & Agnoli, 2012) increase others' well-being but not the donor's material welfare.

Since such nudges do not primarily or directly increase the decision-maker's well-being, they are not justified by libertarian paternalism or are even in contrast with it.

4.3 | Nudging or marketing?

To introduce the third type of nudges, consider a default option to foster saving for retirement and a decoy attempting to steer selection to the most expensive alternative. According to Hansen's narrow definition and Sunstein's broad definition of a nudge, both interventions are nudges, with the former being possibly implemented by the government and the latter by a private company. In fact, they do not forbid any options, they do not significantly alter the decision-maker's incentives, and both exploit his cognitive biases instrumentally. Nevertheless, the two nudges display an important difference: the default benefits the decision-maker, and thus is a pro-self nudge, while the decoy benefits the nudger. Ultimately, the decoy resembles a marketing technique, that is, the attempt to exploit the consumer's biases on the part of the seller to increase her own economic gain.

In effect, many real-world marketing techniques appear to work just like nudges: they preserve the consumer's freedom of choice (e.g., a particular arrangement of items on a shelf or web page) but in fact exploit his cognitive biases (e.g., the automatic renewal of a subscription, which exploits inertia and forgetfulness). For example, as discussed by Congiu and Moscati (2020), Amazon advertises certain products as "Amazon's choice" or "frequently bought," targeting inertia and social conformity, and uses decoys in terms of displaying inferior products together with the product of interest (for an exhaustive overview of marketing techniques from a behavioral-economics perspective, see Foxall, 2016, especially Part II).

Thus, the question is: how can we separate nudges, and specifically those benefitting the nudger, from marketing techniques?

One may deem that libertarian paternalism provides a straightforward criterion to answer this question, namely that a nudge should increase the nudged person's well-being. If we apply this criterion to our initial example, we can conclude that the default to save for retirement is a nudge because it aims to increase the nudged person's well-being, while the decoy to choose the most expensive option is not because it aims to increase the nudger's profit at the expense of the nudgee. Similarly, Amazon's marketing techniques would not be nudges because they primarily aim to increase the firm's profits rather than the consumers' well-being.

However, applying the libertarian paternalistic criterion has two drawbacks. At a general level, it rules out from the nudge toolkit several policies that most scholars consider nudges, namely, "pro-social" nudges (see Section 4.2 above). Secondly, and with specific reference to commercial techniques, these may increase not only the firm's profits but also the consumer's well-being.

For instance, a firm might exploit a salience effect to increase the sales of a smartphone with the best quality/price ratio, thus establishing a win–win transaction where both parties are satisfied. Moreover, a retailer might nudge its customers to buy an expensive refrigerator that is more efficient in terms of energy use, with positive externalities for both the consumers and society, provided that rebound effects are not too large (see, e.g., Blasch et al., 2019). In fact, while the retailer increases its profits, it also increases the customer's welfare in the form of lower electricity bills and society's welfare in terms of lower energy use and pollution. Finally, even in the archetypal cafeteria nudge, the cafeteria manager may want to make healthy food more prominent and accessible to customers also because the profit she makes on that type of food, say organic vegetables, is higher than the profit she makes on chips, candy, and other junk food.

A possible solution to the problem of separating nudges from "mere" marketing techniques consists of incorporating into the nudge definition a specific necessary condition, namely, that an intervention counts as a nudge if it does not increase the nudger's well-being *exclusively*. This criterion makes room for both pro-self and pro-social nudges, which benefit, respectively, the nudged individual and society rather than the nudger, as well as for commercial nudges that increase the firm's welfare as well as the consumers'.

4.4 | Summing up

To conclude our survey of the discussions about the definition of a nudge, we can say that nudges are interventions that attempt to influence people's behavior by exploiting, *at least in part*, their rationality failures, and that do not increase, *exclusively*, the well-being of nudgers. Admittedly, the two clauses—"at least in part" and "not exclusively"—make the definition somehow indefinite, for it is not always straightforward to determine whether they are satisfied. This limitation notwithstanding, on the one hand this definition appears wide enough to capture the variegated policy interventions that, over the last fifteen years or so, have been conceived of or advertised as nudges. On the other hand, the definition seems narrow enough to distinguish nudges from other and more traditional policy interventions, and even from marketing techniques that harm the consumer for the company's exclusive gain.

5 | THE LIBERTARIAN-PATERNALISTIC JUSTIFICATION OF PRO-SELF NUDGES

According to Thaler and Sunstein the political-philosophy view that guides and justifies the use of nudges is libertarian paternalism. This is the approach that "preserves freedom of choice [and in this sense is libertarian] but authorizes both private and public institutions to steer people in directions that will promote their welfare [and in this sense is paternalistic]" (Thaler & Sunstein, 2003, p. 179). Notably, insofar as they are informed by libertarian paternalism, nudges should aim at making decision-makers "better off, *as judged by themselves*" (Thaler & Sunstein, 2009, p. 5; see also Thaler, 2015, pp. 325–326). This latter clause distinguishes the paternalism advocated by Thaler and Sunstein from other forms of paternalism based on religious or moral notions of good that do not depend on the decision-makers' subjective judgment (for a discussion, see Rizzo & Whitman, 2019).

In effect, as argued by Barton and Grüne-Yanoff (2015) and anticipated in Section 4 above, libertarian paternalism provides, at most, a justification for pro-self nudges, while it does not

apply to, or even provides arguments against, pro-social and commercial nudges. For the sake of simplicity, in discussing libertarian paternalistic arguments in support of pro-self nudges, we shall mostly skip the "pro-self" specification and simply talk of "nudges."

For Thaler and Sunstein, libertarian paternalism justifies (pro-self) nudges in three main ways.

First, nudges steer choices in directions that decision-makers themselves would pursue if they were "rational" or, at least, "had complete information, unlimited cognitive abilities, and no lack of self-control" (Sunstein & Thaler, 2003, p. 1162). In other words, nudges "rationalize" the decision-maker's choices, in the sense of aligning his actual choices with those that would be generated by his rational preferences. For instance, if decision-makers recognize the importance of a healthy diet and have in mind the adverse consequences of unhealthy eating, they will gladly choose healthy foods over unhealthy ones. Hence, a food arrangement that helps decision-makers to attain this goal "rationalizes" people's choices and makes them better off.

Second, Thaler and Sunstein (2009, pp. 10, 240, 255) maintain that a choice architecture is always in place, and that people are influenced by it whether someone consciously designs it or not. For example, food items in cafeteria lines have to be arranged in some way, and the arrangement of food influences people's choices irrespective of whether the cafeteria manager deliberately intends to do so. Since an existing choice architecture can steer people's decisions in unintended directions, some of which may even be detrimental for them, Thaler and Sunstein argue in favor of an intentional design that promotes people's well-being.

Third, by their very nature, interventions that fall under the libertarian-paternalistic view preserve the liberty to choose, in the sense that the decision-maker can choose among all existing alternatives and easily avoid or resist the influence attempt if he wants to. An implication is that, since harder forms of behavioral change are already in place and accepted by the public—such as laws, bans, and mandates—then interventions such as nudges, which gently steer behavior instead of forcing it, should be accepted as well (Sunstein, 2018).

6 | ISSUES WITH THE LIBERTARIAN-PATERNALISTIC JUSTIFICATION OF PRO-SELF NUDGES

As has the concept of nudge, the concept of libertarian paternalism has been the subject of intense discussion and criticism.[2] We review the debate on libertarian paternalism with respect to the three justifications it offers for pro-self nudges, namely, (1) the alleged rationalization of choices; (2) the inevitable existence of a choice architecture; and (3) the preservation of liberty and the absence of manipulation.

6.1 | Rationalization of choices

As already mentioned, according to Thaler and Sunstein, libertarian paternalism justifies nudges because they align the decision-maker's behavior with his "true" preferences—those he would have if he had complete information and understanding, unlimited cognitive power, and no lack of self-control. However, this claim poses several problems.

In the first place, the decision-maker may not have "true" preferences at all (Infante et al., 2016; Rizzo & Whitman, 2019; Sugden, 2018). As Infante et al. (2016) and Sugden (2018) argue, the idea of decision-makers having "true" preferences presupposes a model of decision making where an "inner rational agent" with "true" preferences is entrapped in a "outer psychological

shell" responsible for distorting such preferences. However, this model is problematic from both a psychological and philosophical perspective, because the mechanisms occurring in the psychological shell cannot be assumed to simply distort the reasoning process of an inner agent. More plausibly, these psychological mechanisms are an integral part of one's decision-making process or, as Miettinen (2021) puts it, "are exhaustively and tautologically the whole decision-making process". This ultimately implies that it may be impossible to separate "true" preferences from distorted preferences, casting serious doubts about whether "true" preferences really exist.

Second, even assuming that "true" preferences exist, it may be difficult (if not impossible) for the policy maker to know them, that is, she may face a "knowledge problem" (Hands, 2021; Rizzo & Whitman, 2019; Sugden, 2018). For example, in the case of the cafeteria line, it is far from obvious that the true preferences of every customer drive him to eat healthily: some customers may very well be satisfied with an unhealthy diet for specific, personal reasons. For instance, a worker may want to consume sugary food to build up energy sufficient to get through the day. Or, as an extreme case, a terminally ill or pessimistic individual may want to indulge in hedonic pleasures given that his life will end soon. And the cafeteria manager may not be aware of these non-standard but true preferences.

These two criticisms cast doubt on the pretense that nudges can align the decision-maker's behavior with his true preferences, and therefore on the libertarian paternalistic justification of pro-self nudges.

6.2 | Inevitable existence of a choice architecture

A choice architecture may influence the individual's choice in both "good" and "bad" directions, that is, respectively, in directions that increase the well-being of the nudged person or that decrease it. Libertarian paternalism recommends pro-self nudges that steer choice in "good" directions. However, this line of argument raises two issues—one practical, the other ethical.

The practical issue is that a choice architecture, according to some scholars, can be designed without steering the decision in specific directions—that is, in a "neutral" way (Gigerenzer, 2015; Grüne-Yanoff, 2012). As we will discuss in more detail in Section 8, there exists a class of behavioral techniques, termed "boosts," that aim at enhancing the decision-maker's own capabilities without driving him to a specific option. So, for example, suppose that a clinician has to communicate to his patient the risk associated with a surgery, and suppose that natural frequencies ("5 out of 100") are typically more comprehensible than percentages ("5%"). According to Gigerenzer (2015) and boost supporters, the clinician should express such risk in natural frequencies just because they improve his patient's decision making. In his reply to this critique, Sunstein (2015, 2018) contends that many nudges can enhance decision making as boosts do, such as GPS, reminders, and warning labels.

The ethical issue refers to the intentionality of nudging. Some moral philosophers have argued that there is a significant difference between "allowing" something to happen and "causing" something to happen (e.g., Quinn, 1989). Thus, allowing something (good or bad) to occur as a result of an already existing choice architecture is ethically less problematic than causing something (good or bad) to occur as a result of a choice architecture deliberately designed for that purpose. The difference, in fact, lies in the intentionality of the intervention with respect to the behavioral consequences (Barton & Grüne-Yanoff, 2015).

We note, however, that embracing such philosophical arguments about intentionality risks entrapping the whole policy-making enterprise for good, whether behaviorally informed or not.

In fact, under this perspective, *any* intervention in one's environment would be morally wrong because it entails an intentional interference in a certain direction.

6.3 | Liberty preservation and choice manipulation

Some scholars contend that libertarian paternalism, and hence pro-self nudges, are not as libertarian as Thaler and Sunstein suggest.

A first problem concerns the very definition of "liberty preservation" associated with Thaler and Sunstein's idea that a policy is libertarian insofar as it preserves freedom of choice. Thaler and Sunstein (2003, 2009) seem to define liberty preservation as the possibility of selecting any of the available options, without restrictions or constraints. If this view of liberty is adopted, then pro-self nudges are adequately justified by libertarian paternalism because nudges cannot restrict the choice set or impose significant costs on the decision-maker.

However, alternative definitions of liberty preservation may pose challenges to this line of reasoning. For example, Grüne-Yanoff (2012) contends that liberal theorists typically adopt a wider definition of liberty, such as the absence of a "deliberate interference of other human beings within the area in which I could otherwise act" (Berlin, 1969, p. 122). Under this definition, libertarian paternalism would limit liberty. In fact, the specific design of choice architecture to attain some behavioral response (e.g., arranging food items in a cafeteria line to promote healthy eating or setting as the default one's consent to organ donation) is by itself an interference with one's decision. Hence, libertarian paternalism would not be consistent with libertarian views.

A second problem is that liberty preservation, according to some scholars, does not depend only on whether the intervention restricts the choice set or imposes costs on the decision-maker. For example, Saghai (2013) specifies that the fact that a policy intervention does not restrict the choice set is a necessary, but not sufficient, condition for liberty. In order to be truly freedom-preserving, a policy intervention should also trigger cognitive deliberation, so that the decision-maker can "resist" the nudge, and only then can he preserve his liberty. Nonetheless, as argued in Section 2.3, only those nudges which target System 2 of the mind trigger cognitive deliberation. Therefore, only Type 2 nudges preserve liberty in Saghai's sense.

Related to this, a third and final problem discussed in the literature is that an intervention that does not trigger deliberation and that exploits behavioral biases in a covert manner can be manipulative.[3] Manipulation entails the subversion of the decision-maker's autonomy. In effect, Hansen and Jespersen (2013) show that some pro-self nudges, in addition to bypassing cognitive deliberation, are also non-transparent—that is, their existence, means, and intention are not evident to the decision-maker. Hence, these nudges not only limit the decision-maker's liberty but also his decisional autonomy, which further undermines their justifiability. Thus, ultimately, libertarian paternalism may justify only a subset of pro-self nudges, that is, those that trigger cognitive deliberation and are transparent. Among those nudges that do not infringe on one's liberty and autonomy, Sunstein (2015, 2018) often mentions GPS, warning labels, and reminders, which are in fact Type 2 transparent nudges.

7 | JUSTIFYING NUDGES BEYOND LIBERTARIAN PATERNALISM

Because of the limitations of the libertarian paternalist justification, several scholars have put forward alternative justifications for nudges.

7.1 | Justifications for pro-self nudges not based on libertarian paternalism

We argued in Section 6.1 that several critics of libertarian paternalism deem implausible that nudges can align a decision-maker's behavior with his "true" preferences, either because true preferences do not exist or because, if they exist, the policy maker may not know them.

Nonetheless, a policy maker may still find it sensible to nudge the decision-maker in a specific direction irrespective of whether such direction is in accordance with the decision-maker's true preferences. Taking an example from Guala and Mittone (2015), suppose that Ann truly prefers not saving for retirement because currently she is not interested in her standard of living at an older age. Irrespective of Ann's true preference, the sensible course of action for a policy maker could be to nudge Ann to save for retirement, because this would certainly improve the quality of her life at a later stage.

This would be, as Guala and Mittone (2015, p. 386) define it, a "genuinely paternalistic" stance on pro-self nudging, whereby "nudged individuals are better off independently of their preferences." Thus, the use of pro-self nudges that aim to foster physical activity or the adoption of energy plans with lower rates is justified on the ground that it is in the individual's interest to act so because both choices prove better for him in the long run (e.g., better health in the former case, private cost-saving in the latter). Notice that the main difference with hard forms of paternalism would be that nudges remain gentler policies compared to bans or mandates. In fact, instead of nudging the individual to engage in physical activity or adopt cheaper energy plans, a hard paternalist would, for example, make both options mandatory.

7.2 | Justifications for pro-social nudges

In addition to "genuine paternalism," critics of libertarian paternalism argue in favor of a "no harm" (e.g., Oliver, 2013; Oliver, 2017) or "political" (Guala & Mittone, 2015) stance on nudges. The idea is that people's behaviors sometimes generate costs that are not paid by the decision-maker himself (i.e., negative externalities), and nudges may be justified on the ground that they reduce (i.e., internalize) such costs. For example, nudges that foster truthful tax declarations or the adoption of sustainable energy plans can be justified as policies that internalize the negative externalities associated with tax evasion and the deterioration of the planet's resources. Thus, the "no harm" or "political" stance authorizes policy makers to "intervene to protect *other* people from the damage that may be caused by irresponsible individuals [i.e., individuals who create negative externalities through their decisions]. Nudge policies are … for the good of third parties that otherwise are going to be harmed" (Guala & Mittone, 2015, p. 392). Ultimately, the "political" or "no harm" stance is that which justifies the use of pro-social nudges.

7.3 | Justifications based on public acceptance

Finally, a possible justification for nudging, be it pro-self or pro-social, might reside in mere acceptance by the public. Nudge policies often do not require explicit legislation, and therefore for them public acceptance may be more important than it is for taxes and other traditional policies that are approved by parliament or other democratic bodies.[4]

In recent years, several surveys have been carried out to assess people's opinions on nudging, either at the national level, such as in the United States (Gold et al., 2020; Jung & Mellers, 2016) and the United Kingdom (Gold et al., 2020; Reynolds et al., 2019), or across countries, in both Europe (Loibl et al., 2018; Reisch & Sunstein, 2016) and worldwide (Felsen et al., 2013; Hagman et al., 2015; Sunstein et al., 2018, 2019).

At a general level, surveys seem to indicate high acceptance rates for nudges (Reisch & Sunstein, 2016; Sunstein, 2017b; Sunstein et al., 2018). However, public acceptance varies depending on factors such as the cognitive system targeted by the nudge (i.e., Type 1 vs. Type 2 nudges), policy domain (e.g., obesity vs. organ donation), country, and targeted well-being (i.e., pro-self vs. pro-social nudges).

With respect to the targeted cognitive system, Jung and Mellers (2016) show that US citizens are more favorably inclined towards Type 2 transparent nudges, and that more covert interventions are seen as manipulative; this appears to be a recurrent finding in many surveys (Felsen et al., 2013; Gold et al., 2020; Kasdan, 2020; Lin et al., 2017; Sunstein, 2016). With respect to policy domain and country, people approve of the use of nudges to promote health and safety (Reynolds et al., 2019; Sunstein et al., 2018). Still, there is the important exception of default consent in organ donation, for which the study by Hagman et al. (2015) finds high opposition rates in both samples considered (Sweden and US citizens). Finally, acceptance rates vary based on whether nudges are pro-self or pro-social: surveys show that the public's attitudes are more favorable in the case of pro-social nudges, such as those addressing tax evasion, obesity, and tobacco and alcohol consumption (Hagman et al., 2015; Reynolds et al., 2019; Sunstein, 2015; see also Sunstein, 2017b).

In sum, the public's opinions on nudges may be used as a yardstick to validate them and so legitimize their use.

8 | BOOSTS VERSUS NUDGES

In partial opposition to nudges, Grüne-Yanoff, Hertwig, and co-authors have recently proposed alternative behavioral policies they call "boosts" (Grüne-Yanoff, 2018; Grüne-Yanoff et al., 2018; Grüne-Yanoff & Hertwig, 2016; Hertwig & Grüne-Yanoff, 2017;). We think it proper to briefly illustrate what boosts are and discuss their relationship with nudges.

Advocates of both nudges and boosts agree that human decision making does not conform to the neoclassical notion of rationality and is often defective. Both seek to change defective behaviors without prohibiting options or substantially changing incentives. However, while nudges attempt to change people's behavior by exploiting their rationality failures, boosts attempt to help people to overcome such failures. More precisely, Grüne-Yanoff and Hertwig (2016, p. 156) define boosts as behavioral policies that have "the goal of empowering people by expanding (boosting) their competences and thus helping them to reach their objectives." In particular, boosts can either "change the environment in which decisions are made [or] extend the repertoire of decision-making strategies, skills, and knowledge, or do both" (p. 156).

An example of a boost changing the decision environment is the one already mentioned in Section 6.2, namely, the presentation of risk information in a numerical format that enhances its comprehension. The studies of Gigerenzer and Hoffrage (1995) and of Hoffrage et al. (2000) have in fact shown that people typically have difficulties in solving a problem whose risk information is conveyed in terms of probabilities expressed as a percentage (e.g., "5%"), but when risk information is restated in natural frequencies (e.g., "5 out of 100"), the number of correct responses increases dramatically. An example of a boost that extends the repertoire of decision-making skills

is a diagnostic technique that allows performing quick medical diagnoses based on few, selected physical cues. For instance, Fischer et al. (2002) and Marewski and Gigerenzer (2012) argue that an accurate prediction tool to treat pneumonia in young children can be designed as a simple decision tree based on two simple cues—duration of fever and child's age. Similar techniques can help laypeople to recognize, for instance, the signs of a heart attack or stroke, and seek timely help (Grüne-Yanoff & Hertwig, 2016).

One significant advantage of boosts over nudges is that the former can be justified without the shortcomings of the libertarian paternalist justification of (pro-self) nudges (see Section 6). In the first place, while nudges are based on the implicit assumption that the policy maker knows the true preferences and goals of decision-makers (see Section 6.1), boosts aim at expanding the decision-maker's ability to reach his own objectives, "without making undue assumptions about what those objectives are" (Grüne-Yanoff & Hertwig, 2016, p. 156). Moreover, when a boost alters the decision context, it does so to make it more comprehensible or accessible, such as when risk information is restated to favor correct statistical inferences. This implies that boosts allow policy makers to design a choice architecture in a "neutral" way, that is, one that avoids steering a decision in specific directions (see Section 6.2), and one that does not infringe on the decision-maker's autonomy (see Section 6.3).

Despite these differences, in practice the distinction between nudges and boosts is not always clear-cut. Although Grüne-Yanoff and his co-authors have attempted to lay down a series of separating criteria, boosts still share remarkable similarities with a relevant subset of nudges, namely, Type 2 transparent nudges (see Sections 2.3 and 6.3). In fact, nudges of this type are capable of educating the decision-maker and empowering his decision in a similar fashion as boosts do. Accordingly, the "look right, look left" sign in a London street mentioned in Section 2.3 can be seen both as a Type 2 transparent nudge and as a boost that attempts to lead the pedestrian to adopt safer behavior. Likewise, a boost that teaches the decision-maker to "think" in natural frequencies when confronted with risky choices can be thought of as a Type 2 transparent nudge, since it targets the decision-maker's reflective system and has the evident purpose of augmenting his statistical literacy.

Besides, while the boost program entails a promising collection of behavioral techniques, the effectiveness and breadth of applications of boosts seem more limited than that of nudges. For example, boosting the decision-maker's ability to compute the optimal saving rate does not guarantee that the decision-maker will actually start saving for retirement, while a nudge such as the default enrollment in a saving plan might go a longer way. Likewise, boosting one's understanding of the risk of developing cardiovascular diseases will not necessarily persuade an obese individual to start following a diet, while a nudge might perhaps circumvent his lack of self-control.

In sum, we deem that in some situations boosts actually enrich the toolkit of behavioral policy makers, but we see boost policies more as a complement to than a substitute for nudge policies.

9 | EFFECTIVENESS OF NUDGING

Research demonstrates that nudging can be effective in various domains, such as promoting healthy behaviors, saving for retirement, adopting pro-environmental conduct, and fostering tax compliance.[5]

Despite the many cases of success, there might be situations where a nudge does not work as expected (Sunstein, 2017a). Even the same type of nudge may work in one scenario and fail

in another. As an example, consider the use of defaults—one of the most effective and robust nudges. The first contributions to document their effectiveness are those by Johnson and Goldstein (2003, 2004), who use an online experiment and data from various countries to show that contexts where the default is to be organ donors see dramatically higher donation rates compared to contexts with the opposite default. The effectiveness of defaults has been subsequently proven in numerous studies and in the most disparate sectors (e.g., Bruns et al., 2018; Cronqvist et al., 2018; Jachimowicz et al., 2019; Venema et al., 2018; Wynes et al., 2018).

However, even robust nudges such as defaults can sometimes fail. For example, in a field experiment by Bronchetti et al. (2013), low-income tax filers could either opt-in to allocate part of their tax refunds to US Saving Bonds (control) or opt-out from a default program that performed this allocation automatically (treatment). Results show that saving behavior did not change across treatments (i.e., conditional on the default allocation) and that this might have occurred, as the authors note, because this segment of the population had strong intentions to spend their refunds instead of saving them.

The assessment of the empirical effectiveness of nudges becomes then a crucial question. In what follows, we address it by focusing on the three specific facets of the effectiveness of nudging that we have referred to or hinted at over the course of the article. Specifically, we proceed to discuss (1) whether making a nudge transparent undermines its effectiveness, (2) whether the targeted cognitive system can predict the persistence over time of the behavioral change, (3) whether nudges work better or worse than standard policies, and (4) whether nudges crowd out or complement standard policies.

9.1 | Do nudges work best if non-transparent?

In Section 2.3, it was argued that nudges are transparent insofar as their existence, workings, and intention can be easily reconstructed by the decision-maker. Nonetheless, for several nudges it might be difficult to detect their existence in the choice architecture and reconstruct their workings and intention. In addition to the ethical concerns associated with their use, as already discussed, some scholars have also speculated that nudges that are non-transparent may even work better than transparent nudges. With transparent nudges, in fact, the decision-maker would see through the "trick" and resist the nudging attempt (see, e.g., Bovens, 2009; Grüne-Yanoff, 2012). But is this always the case? Although limited to but a few contributions, research shows that making a default option transparent does not hamper its effectiveness (Bruns et al., 2018; Kroese et al., 2016; Loewenstein et al., 2015; Marchiori et al., 2017; Paunov et al., 2019; Steffel et al., 2016).

For example, Bruns et al. (2018) ran a laboratory experiment where participants were endowed with €10 and asked to donate a portion of it to a "climate protection fund." In each treatment, the participants could either donate a predetermined amount of €8 or manually insert the value they preferred. In addition, they were exposed either to a message explaining that the "preselected value might have an influence on your decision," or that "the preselected value is meant to encourage higher contributions for the climate protection fund," or a combination of both (p. 45). In other words, the experimenters made explicit either the influence attempt, or the purpose of the nudge, or both. Results showed no statistically significant differences with respect to the final contribution across treatments. Thus, this experiment suggests that transparency does not hinder the effectiveness of a default.

In some cases, disclosing a default may even enhance its effectiveness. For example, Paunov et al. (2019) asked participants to imagine they were enrolled in a master's program and that a set of courses had been preselected for them. The treatment where the default was made transparent (through a brief description of the default effect) displayed higher compliance rates with the default. Besides, participants in the non-transparent condition declared that they felt the university did not approach them in a sincere way. These findings suggest that disclosing the default may have an impact on the perceived trustworthiness of the choice architect and, in turn, on compliance rates.

However, while transparency seems not to diminish a default's effectiveness and in some cases may even enhance it, further research is needed to understand to what extent this also applies to other types of nudge.

9.2 | Do nudges work in the long run?

The second question concerns whether the behavioral change generated by a nudge intervention is persistent over time (Marteau et al., 2011, 2017). There are two channels through which the long-term effect of a nudge may be sustained.

First, a nudge may be implemented in the choice architecture only for a limited time and the behavioral change may somehow persist in the nudged subjects, for example because the nudge created a novel habit (Bernedo et al., 2014; Brandon et al., 2017; Burger & Shelton, 2011; Giné et al., 2010). For instance, Burger and Shelton (2011) show that people exposed to signs informing them that most of their peers prefer stairs over the elevator not only adopt the indicated behavior themselves but also adopt it for some time after the signs are removed.

Alternatively, a nudge may be permanently present in the choice architecture and its effect may be renewed each time the decision-maker interacts with it. For instance, Venema et al. (2018) study the long-term effect on a company's employees of setting their "sit-stand desks" (SSDs) on the "stand" position rather than the "sit" position. That is, by default, employees had to work standing instead of sitting. Results show that after two months around 70% of employees were still working standing, that is, complying with the default. Similarly, Cronqvist et al. (2018) report that a default saving plan implemented in Sweden in 2000 (termed the "Premium Pension Plan") was still effective in raising saving rates after nearly two decades (2000–2016).

At present, it is still not clear when the effect of a nudge is likely to persist over time and when it is not. One aspect that has been explored is whether the long-term effectiveness of a nudge may depend on being a Type 1 or a Type 2 nudge (Lin et al., 2017).

For instance, Type 1 nudges—that is, those targeting automatic behavior—might work in the long run because they subconsciously create a habit with respect to some action or simply because they override immediate desires, for instance to indulge in unhealthy eating or smoking. And in fact, the contributions discussed above—the SSDs default and the Swedish saving plan—are examples of Type 1 nudges. However, there is not sufficient evidence to confidently state whether or when Type 1 nudges can or cannot produce sustained behavioral changes.

A similar analysis can be applied to Type 2 nudges. These nudges may in principle work in the long run because they educate or persuade the decision-maker at a conscious level of the desirability of some option. An example is the use of messages that attempt to alter the attitude with respect to a certain behavior, such as that towards the use of stairs (as in Burger & Shelton, 2011). Similarly, Allcott and Rogers (2014) document that the behavioral effect of sending energy consumption reports containing a comparison with the energy consumption of one's peers, albeit

decreasing, was persistent over a period of two years. However, while many Type 2 nudges work in the long run (see, e.g., Bernedo et al., 2014; Kuhfuss et al., 2016; Wakefield et al., 2015; White et al., 2015), research points also to the opposite (Foxcroft et al., 2015). In their comprehensive review of 70 studies, Foxcroft and co-authors found that correcting the misperception of social norms with respect to drinking habits among college students had little to no effect in the long run.

Therefore, while the idea of predicting the long-term effectiveness through the "Type 1–Type 2" lens may be promising, this route still needs further assessment. For example, some additional variables have to be taken into account, such as the credibility of the information source, the specificity of the targeted behavior, and the strength of one's preferences, intentions, and habits with respect to the behavior to change.

9.3 | Do nudges work worse than standard policies?

The third question concerns the issue of how effective nudges are in comparison to standard policies. Some researchers have argued that standard policies are more effective than behavioral interventions because the former influence more strongly and more directly the decision-maker's costs and incentives (see e.g. Baldwin, 2014; Goodwin, 2012; Marteau et al., 2011; Selinger & Whyte, 2012;). While this sometimes may be the case, the converse also happens to be true. In fact, there is a flourishing literature demonstrating that nudges can be more effective (and cost-effective) than standard policy interventions (Benartzi et al., 2017; Chetty et al., 2014; Sunstein, 2018).

For example, Benartzi et al. (2017) calculate the "impact per dollar" of both nudges and standard policies to enhance saving for retirement, students' enrolment in college, energy conservation, and influenza vaccination. Specifically, they evaluate how much every dollar spent on a particular intervention, adjusted for inflation, affected a common dependent variable for that sector (e.g., money allocated to a saving fund). Overall, nudge interventions dramatically outperformed standard policies. For instance, the best performing nudge in the domain of saving for retirement was an "active decision" nudge, which invited new hires in a company to choose their contribution to a retirement fund (Carroll et al., 2009). The best performing standard policy in the same domain was a "retirement savings information" policy, whereby a university's employees were paid $20 to attend a "benefits fair" to receive information about retirement savings plans (Duflo & Saez, 2003). The best performing nudge achieved a return of $100 for each dollar invested, while the best performing standard policy only achieved a return of approximately $15.

However, in some cases the difference between a nudge and standard policy is less dramatic. For example, Benartzi et al. (2017) have compared a "planning-prompt" nudge, where workers were prompted to write down details about when they planned to obtain vaccinations (Milkman et al., 2011), with an educational campaign that also offered free workplace vaccination (Kimura et al., 2007). The comparison shows that, for every $100 dollars spent, the "planning-prompt" nudge achieved a return of approximately 13 adults vaccinated versus nine for the educational campaign. Considering nonetheless that at least some hundreds of dollars would be spent on an intervention of this kind, the resulting difference would accordingly increase in magnitude; for example, investing $100,000 in the educational campaign would mean 4000 fewer people vaccinated compared to the planning-prompt nudge.

In sum, nudges seem not be overshadowed by standard policies, and in some emblematic cases they perform considerably better, at least under a cost-effectiveness perspective.

9.4 | Do nudges crowd out or complement standard policies?

The relationship between nudges and standard policies is not restricted to the issue of their comparative effectiveness, but involves also their possible substitutability or complementarity. Nudges in fact can be perceived as substitutes for standard policies and thus may crowd out the latter.

For instance, Werfel (2017) found that green nudge policies encouraging Japanese households to reduce their energy consumption, e.g., by using fans rather than air conditioning, decreased the households' support for standard, pro-environment policies such as carbon taxes. Hagmann et al. (2019) conducted a series of experiments in which subjects were asked to decide whether or not to implement a carbon tax policy or a green nudge policy, and found that introducing the green nudge reduced the subject's support for the carbon tax.

The authors of the two studies envisage a similar psychological mechanism generating this "nudging-out effect," as it is sometimes called: Individuals perceive green nudges and carbon taxes as substitute means to achieve the goal of climate change mitigation and, when available, they opt for the least burdensome means, namely nudges. This psychological explanation is supported by a further experimental finding of Hagmann et al. (2019), namely that the nudging-out effect disappears when the nudge and the standard policy relate to different policy goals. More specifically, Hagmann et al. (2019) found that a nudge to promote retirement saving diminishes support for social security taxation but does not affect support for carbon taxes.

As Hagmann et al. (2019) notice, the problem with the nudging-out effect is that, if standard policies are in fact more effective than nudge policies, introducing nudges may create "the false hope that problems can be tackled without imposing considerable costs" (p. 485).

Rather than operating as substitutes, in some situations nudges and standard policies can complement each other. According to some scholars (e.g., Banerjee & John, 2021; Bhargava & Loewenstein, 2015; Loewenstein & Chater, 2017), in fact, nudges can be used to enhance the design and effectiveness of standard policies, and vice versa, resulting in what are termed "hybrid" (Loewenstein & Chater, 2017) or "nudge-plus" policies (Banerjee & John, 2021).

For example, a standard policy to increase saving for retirement could make it mandatory for citizens to save a minimum share of their monthly salary, say 9%. However, because of inertia, some citizens may fail to increase their saving rate beyond 9% even if they have the financial resources to do so, and therefore they may still end up saving too little for their retirement. On the nudging side, a policy like "Save More Tomorrow" (Thaler & Benartzi, 2004) automatically enrolls citizens in a saving plan that starts with low saving rates (3.5%) and then, to avoid inertia, increases the existing rate at each salary rise by, say, 2%, until a certain is upper bound is reached. Since "Save More Tomorrow" is a nudge policy, citizens are free to opt-out at any moment. As a consequence, some citizens may still fail to adequately save for their retirement. A hybrid policy could incorporate the "increasing-rates" mechanism into a mandatory saving plan. This policy would ensure higher savings for all citizens, because citizens who would normally not save are now forced to do so, and those who would not increase their saving rates because of inertia are now nudged to do so.

10 | CONCLUSIONS

In this article we have reviewed the nudge literature by focusing on three key topics—the definition of a nudge, the justification of nudges, and their effectiveness. Among other things, we

have discussed the intricate relationship between nudges and marketing techniques, the possibility of justifying nudges beyond libertarian paternalism, and the effectiveness of nudges as compared to the effectiveness of standard policies. With respect to each of those three topics, we have called attention to some open issues that deserve further investigation. Hopefully, this review will "nudge" future researchers to investigate these still open questions.

ACKNOWLEDGEMENTS

We are grateful to two anonymous referees for many insightful suggestions, and to Paolo Crosetto, Francesco Figari, Giulia Galimberti, and Matteo Ploner for helpful comments on previous versions of the manuscript. Moscati thanks Luca Gandullia of Genoa University, and the Camogli Center for Economic Studies for kind hospitality when working on this article. Any errors are ours.

ORCID

Luca Congiu https://orcid.org/0000-0002-6324-7460
Ivan Moscati https://orcid.org/0000-0002-8995-0015

NOTES

[1] See Camerer et al. (2003), Sunstein and Thaler (2003), Bernheim and Rangel (2007), Kőszegi and Rabin (2007), Loewenstein and Haisley (2008), Bernheim and Rangel (2009), and Dhami (2016, chapter 22). For reconstructions of the rise of behavioral economics and behavioral welfare economics richer than the one sketched in this section, see Schoemaker (1982), Rabin (1998), Starmer (2000), Frederick et al. (2002), Etner et al. (2012), McQuillin and Sugden (2012), Angner and Loewenstein (2012), Heukelom (2014), and Oliver (2017).

[2] See Berg and Gigerenzer (2010), Hausman (2012), Rebonato (2012, 2014), Heilmann (2014), Barton and Grüne-Yanoff (2015), Gigerenzer (2015), Guala and Mittone (2015), Le Grand and New (2015), Whitman and Rizzo (2015), Hausman (2016, 2018), Sugden (2018), Rizzo and Whitman (2019), Scoccia (2019), and Hands (2020, 2021).

[3] See Bovens (2009), Grüne-Yanoff (2012), Rebonato (2012), Hansen and Jespersen (2013), Baldwin (2014), and Rebonato (2014).

[4] We thank one of the anonymous referees for calling our attention to this point.

[5] For some examples of nudges promoting health, see Behavioural Insights Team (2016), Lin et al. (2017), Cadario and Chandon (2020), and Gold et al. (2020); for nudges fostering saving for retirement, see Benartzi et al. (2017), Glazebrook et al. (2017), and Cronqvist et al. (2018); for nudges promoting pro-environment conduct, see Schubert (2017), Wynes et al. (2018), Behavioural Insights Team (2020), Buckley (2020), and Daminato et al. (2021); for nudges increasing tax compliance, see Behavioural Insights Team (2016), Kettle et al. (2016), and Antinyan and Asatryan (2020).

REFERENCES

Allais, M. (1953) Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica*, *21*(4), 503–546.

Allcott, H. (2011) Social norms and energy conservation. *Journal of Public Economics*, *95*(9–10), 1082–1095.

Allcott, H., & Rogers, T. (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, *104*(10), 3003–3037.

Angner, E., & Loewenstein, G. (2012) Behavioral economics. In U. Mäki (Ed.), *Handbook of the philosophy of science: Philosophy of economics* (pp. 641–689). Elsevier.

Antinyan, A., & Asatryan, Z. (2020) *Nudging for tax compliance: A meta-analysis*. CESifo. Working Paper No. 8500.

Ariely, D. (2009) *Predictably irrational: The hidden forces that shape our decisions*. Harper Collins.

Baldwin, R. (2014) From regulation to behaviour change: Giving nudge the third degree. *Modern Law Review*, *77*(6), 831–857.

Banerjee, S., & John, P. (2021) Nudge plus: Incorporating reflection into behavioural public policy. *Behavioural Public Policy*, 1–16. https://doi.org/10.1017/bpp.2021.6

Barton, A., & Grüne-Yanoff, T. (2015) From libertarian paternalism to nudging—and beyond. *Review of Philosophy and Psychology*, *6*(3), 341–359.

Behavioural Insights Team (2016) *The behavioural insights team's update report: 2015-16*. https://www.bi.team/publications/the-behavioural-insights-teams-update-report-2015-16/

Behavioural Insights Team (2020) *The little book of green nudges*. https://www.bi.team/publications/the-little-book-of-green-nudges/

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

Berg, N., & Gigerenzer, G. (2010) As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, *18*, 133–166.

Berlin, I. (1969) Two concepts of liberty. In I. Berlin (Ed.), *Four essays on liberty* (pp. 118–172). Oxford University Press.

Bernedo, M., Ferraro, P. J., & Price, M. (2014) The persistent impacts of norm-based messaging and their implications for water conservation. *Journal of Consumer Policy*, *37*(3), 437–452.

Bernheim, B. D., & Rangel, A. (2007) Toward choice-theoretic foundations for behavioral welfare economics. *American Economic Review*, *97*(2), 464–470.

Bernheim, B. D., & Rangel, A. (2009) Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, *124*(1), 51–104.

Berthet, V., & Ouvrard, B. (2019) Nudge: Towards a consensus view? *Psychology and Cognitive Science*, *5*, 1–5.

Bhargava, S., & Loewenstein, G. (2015) Behavioral economics and public policy 102: Beyond nudging. *American Economic Review*, *105*(5), 396–401.

Blasch, J., Filippini, M., & Kumar, N. (2019) Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource and Energy Economics*, *56*, 39–58.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013) Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803–843.

Bovens, L. (2009) The ethics of nudge. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 207–219). Springer.

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2017) *Do the effects of social nudges persist? Theory and evidence from 38 natural field experiments*. National Bureau of Economic Research. Working Paper 23277.

Bronchetti, E. T., Dee, T. S., Huffman, D. B., & Magenheim, E. (2013) When a nudge isn't enough: Defaults and saving among low-income tax filers. *National Tax Journal*, *66*(3), 609–634.

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., & Rahali, B. (2018) Can nudges be transparent and yet effective? *Journal of Economic Psychology*, *65*, 41–59.

Buckley, P. (2020) Prices, information and nudges for residential electricity conservation: A meta-analysis. *Ecological Economics*, *172*, 106635.

Burger, J. M., & Shelton, M. (2011) Changing everyday health behaviors through descriptive norm manipulations. *Social Influence*, *6*(2), 69–77.

Cadario, R., & Chandon, P. (2020) Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science*, *39*(3), 465–486.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003) Regulation for conservatives: Behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review*, *151*(3), 1211–1254.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2009) Optimal defaults and active decisions. *Quarterly Journal of Economics*, *124*(4), 1639–1674.

Chernoff, H. (1954) Rational selection of decision functions. *Econometrica*, *4*(22), 422–443.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014) Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Quarterly Journal of Economics*, *129*(3), 1141–1219.

Chetty, R., Looney, A., & Kroft, K. (2009) Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145–1177.

Cialdini, R. B., & Goldstein, N. J. (2004) Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621.

Congiu, L., & Moscati, I. (2020) Message and Environment: A framework for nudges and choice architecture. *Behavioural Public Policy*, *4*(1), 71–87.

Cronqvist, H., Thaler, R. H., & Yu, F. (2018) When nudges are forever: Inertia in the swedish premium pension plan. *AEA Papers and Proceedings*, *108*, 153–158.

Daminato, C., Diaz-Farina, E., Filippini, M., & Padrón-Fumero, N. (2021) The impact of smart meters on residential water consumption: Evidence from a natural experiment in the Canary Islands. *Resource and Energy Economics*, *64*, 101221.

Dhami, S. (2016) *The foundations of behavioral economic analysis*. Oxford University Press.

Duflo, E., & Saez, E. (2003) The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, *118*(3), 815–842.

Ellsberg, D. (1961) Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*(4), 643–669.

Etner, J., Jeleva, M., & Tallon, J.-M. (2012) Decision theory under ambiguity. *Journal of Economic Surveys*, *26*(2), 234–270.

Felsen, G., Castelo, N., & Reiner, P. B. (2013) Decisional enhancement and autonomy: Public attitudes towards overt and covert nudges. *Judgment and Decision Making*, *8*(3), 202–213.

Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., Altwegg, M., & Nadal, D. (2002) Use of simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Archives of Pediatrics and Adolescent Medicine*, *156*(10), 1005–1008.

Foxall, G. R. (Ed.) (2016) *The Routledge companion to consumer behavior analysis*. Routledge.

Foxcroft, D. R., Moreira, M. T., Santimano, N. M. A., & Smith, L. A. (2015) Social norms information for alcohol misuse in university and college students. *Cochrane Database of Systematic Reviews*, *1*. https://doi.org/10.1002/14651858.CD006748.PUB3

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002) Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Gaudeul, A., & Crosetto, P. (2019) *Fast then slow: A choice process explanation for the attraction effect*. Grenoble Applied Economics Laboratory (GAEL). Working Paper 2019-06.

Gigerenzer, G. (2015) On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology*, *6*(3), 361–383.

Gigerenzer, G., & Hoffrage, U. (1995) How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*(4), 684–704.

Giné, X., Karlan, D., & Zinman, J. (2010) Put your money where your butt is: A commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, *2*(4), 213–235.

Glazebrook, K., Larkin, C., & Costa, E. (2017) *Improving engagement with pension decisions: the results from three randomised controlled trials*. https://www.bi.team/wp-content/uploads/2017/10/Pension-wise-trials.pdf

Gold, N., Lin, Y., Ashcroft, R., & Osman, M. (2020) 'Better off, as judged by themselves': Do people support nudges as a method to change their own behavior? *Behavioural Public Policy*, 1-30. https://doi.org/10.1017/bpp.2020.6

Goodwin, T. (2012) Why we should reject 'Nudge'. *Politics*, *32*(2), 85–92.

Grüne-Yanoff, T. (2012) Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, *38*(4), 635–645.

Grüne-Yanoff, T. (2018) Boosts vs. nudges from a welfarist perspective. *Revue d'Économie Politique*, *128*(2), 209–224.

Grüne-Yanoff, T., & Hertwig, R. (2016) Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, *26*(1–2), 149–183.

Grüne-Yanoff, T., Marchionni, C., & Feufel, M. A. (2018) Toward a framework for selecting behavioural policies: How to choose between boosts and nudges. *Economics and Philosophy*, *34*, 243–266.

Guala, F., & Mittone, L. (2015) A political justification of nudging. *Review of Philosophy and Psychology*, *6*(3), 385–395.

Hagman, W., Andersson, D., Västfjäll, D., & Tinghög, G. (2015) Public views on policies involving nudges. *Review of Philosophy and Psychology*, *6*(3), 439–453.

Hagmann, D., Ho, E. H., & Loewenstein, G. (2019) Nudging out support for a carbon tax. *Nature Climate Change*, *9*(6), 484–489.

Handa, J. (1977) Risk, probabilities, and a new theory of cardinal utility. *Journal of Political Economy*, *85*(1), 97–122.

Hands, D. W. (2020) Libertarian paternalism: Taking Econs seriously. *International Review of Economics*, *67*(4), 1–23.

Hands, D. W. (2021) Libertarian paternalism: Making rational fools. *Review of Behavioral Economics*, Forthcoming.
Hansen, P. G. (2016) The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation*, *7*(1), 155–174.
Hansen, P. G., & Jespersen, A. M. (2013) Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation*, *4*(1), 3–28.
Hausman, D. M. (2012) *Preference, value, choice, and welfare*. Cambridge University Press.
Hausman, D. M. (2016) On the Econ within. *Journal of Economic Methodology*, *23*(1), 26–32.
Hausman, D. M. (2018) Behavioural economics and paternalism. *Economics and Philosophy*, *34*(1), 53–66.
Hausman, D. M., & Welch, B. (2010) Debate: To nudge or not to nudge. *Journal of Political Philosophy*, *18*(1), 123–136.
Heilmann, C. (2014) Success conditions for nudges: A methodological critique of libertarian paternalism. *European Journal for Philosophy of Science*, *4*(1), 75–94.
Hertwig, R., & Grüne-Yanoff, T. (2017) Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973–986.
Heukelom, F. (2014) *Behavioral economics: A history*. Cambridge University Press.
Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000) Communicating statistical information. *Science*, *290*(5500), 2261–2262.
Homonoff, T. A. (2018) Can small incentives have large effects? The impact of taxes versus bonuses on disposable bag use. *American Economic Journal: Economic Policy*, *10*(4), 177–210.
Infante, G., Lecouteux, G., & Sugden, R. (2016) Preference purification and the inner rational agent: A critique of the conventional wisdom of behavioural welfare economics. *Journal of Economic Methodology*, *23*(1), 1–25.
Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019) When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, *3*(2), 159–186.
John, P., & Blume, T. (2018) How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a central London local authority. *Journal of Behavioral Public Administration*, *1*(1), 1–11.
Johnson, E. J., & Goldstein, D. G. (2003) Do defaults save lives? *Science*, *302*(5649), 1338–1339.
Johnson, E. J., & Goldstein, D. G. (2004) Defaults and donation decisions. *Transplantation*, *78*(12), 1713–1716.
Jung, J. Y., & Mellers, B. A. (2016) American attitudes toward nudges. *Judgment and Decision Making*, *11*(1), 62–74.
Kahneman, D. (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux.
Kahneman, D., & Tversky, A. (1979) Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*(2), 263–292.
Kasdan, D. O. (2020) Do Koreans like being nudged? Survey evidence for the contextuality of behavioral public policy. *Rationality and Society*, *32*(3), 313–333.
Kettle, S., Hernandez, M., Ruda, S., & Sanders, M. (2016) *Behavioral interventions in tax compliance: Evidence from Guatemala*. World Bank Policy Research, Working Paper No. 7690.
Kimura, A. C., Nguyen, C. N., Higa, J. I., Hurwitz, E. L., & Vugia, D. J. (2007) The effectiveness of vaccine day and educational interventions on influenza vaccine coverage among health care workers at long-term care facilities. *American Journal of Public Health*, *97*(4), 684–690.
Kivetz, R., Netzer, O., & Srinivasan, V. (2004) Extending compromise effect models to complex buying situations and other context effects. *Journal of Marketing Research*, *41*(3), 262–268.
Kőszegi, B., & Rabin, M. (2007) Mistakes in choice-based welfare analysis. *American Economic Review*, *97*(2), 477–481.
Kroese, F. M., Marchiori, D. R., & de Ridder, D. T. (2016) Nudging healthy food choices: A field experiment at the train station. *Journal of Public Health*, *38*(2), e133–e137.
Kuhfuss, L., Préget, R., Thoyer, S., Hanley, N., Le Coent, P., & Désolé, M. (2016) Nudges, social norms, and permanence in agri-environmental schemes. *Land Economics*, *92*(4), 641–655.
Le Grand, J., & New, B. (2015) *Government paternalism: Nanny state or helpful friend?*. Princeton University Press.
Lin, Y., Osman, M., & Ashcroft, R. (2017) Nudge: Concept, effectiveness, and ethics. *Basic and Applied Social Psychology*, *39*(6), 293–306.
Loewenstein, G., & Chater, N. (2017) Putting nudges in perspective. *Behavioural Public Policy*, *1*(1), 26–53.
Loewenstein, G., & Haisley, E. (2008) The economist as therapist: Methodological ramifications of "light" paternalism. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics* (pp. 210–245). Oxford University Press.

Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015) Warning: You are about to be nudged. *Behavioral Science and Policy*, *1*(1), 35–42.

Loibl, C., Sunstein, C. R., Rauber, J., & Reisch, L. A. (2018) Which Europeans like nudges? Approval and controversy in four European countries. *Journal of Consumer Affairs*, *52*(3), 655–688.

Loomes, G., & Sugden, R. (1982) Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, *92*(368), 805–824.

Marchiori, D. R., Adriaanse, M. A., & de Ridder, D. T. (2017) Unresolved questions in nudging research: Putting the psychology back in nudging. *Social and Personality Psychology Compass*, *11*(1), e12297.

Marewski, J. N., & Gigerenzer, G. (2012) Heuristic decision making in medicine. *Dialogues in Clinical Neuroscience*, *14*(1), 77–89.

Marteau, T. M., Ogilvie, D., Roland, M., Suhrcke, M., & Kelly, M. P. (2011) Judging nudging: Can nudging improve population health? *British Medical Journal*, *342*, d228.

McQuillin, B., & Sugden, R. (2012) Reconciling normative and behavioural economics: The problems to be solved. *Social Choice and Welfare*, *38*(4), 553–567.

Miettinen, T. (2021) Review on Robert Sugden's "the community of advantage: A behavioural economist's defence of the market". Oxford, July 17th 2018. 352 p. ISBN: 9780198825142. *Journal of Behavioral and Experimental Economics*, *92*. https://doi.org/10.1016/j.socec.2021.101683

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

Mongin, P., & Cozic, M. (2018) Rethinking nudge: Not one but three concepts. *Behavioural Public Policy*, *2*(1), 107–124.

Nagatsu, M. (2015) Social nudges: Their mechanisms and justification. *Review of Philosophy and Psychology*, *6*(3), 481–494.

Oliver, A. (2013) From nudging to budging: Using behavioural economics to inform public sector policy. *Journal of Social Policy*, *42*(4), 685–700.

Oliver, A. (2017) *The origins of behavioural public policy*. Cambridge University Press.

Paunov, Y., Wänke, M., & Vogel, T. (2019) Transparency effects on policy compliance: Disclosing how defaults work can enhance their effectiveness. *Behavioural Public Policy*, *3*(2), 187–208.

Pittarello, A., Caserotti, M., & Rubaltelli, E. (2019) 'Three is better than two': Increasing donations with the attraction effect. *British Journal of Psychology*, *111*(4), 805–822.

Quinn, W. S. (1989) Actions, intentions, and consequences: The doctrine of doing and allowing. *Philosophical Review*, *98*(3), 287–312.

Rabin, M. (1998) Psychology and economics. *Journal of Economic Literature*, *36*(1), 11–46.

Rebonato, R. (2012) *Taking liberties: A critical examination of libertarian paternalism*. Palgrave Macmillan.

Rebonato, R. (2014) A critical assessment of libertarian paternalism. *Journal of Consumer Policy*, *37*(3), 357–396.

Reisch, L. A., & Sunstein, C. R. (2016) Do Europeans like nudges? *Judgment and Decision Making*, *11*(4), 310–325.

Reynolds, J., Archer, S., Pilling, M., Kenny, M., Hollands, G. J., & Marteau, T. (2019) Public acceptability of nudging and taxing to reduce consumption of alcohol, tobacco, and food: A population-based survey experiment. *Social Science and Medicine*, *236*, 112395.

Rizzo, M. J., & Whitman, G. (2019) *Escaping paternalism: Rationality, behavioral economics, and public policy*. Cambridge University Press.

Rubaltelli, E., & Agnoli, S. (2012) The emotional cost of charitable donations. *Cognition and Emotion*, *26*(5), 769–785.

Saghai, Y. (2013) Salvaging the concept of nudge. *Journal of Medical Ethics*, *39*(8), 487–493.

Schoemaker, P. J. H. (1982) The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, *20*(2), 529–563.

Schubert, C. (2017) Green nudges: Do they work? Are they ethical? *Ecological Economics*, *132*, 329–342.

Scoccia, D. (2019) Paternalisms and nudges. *Economics and Philosophy*, *35*(1), 79–102.

Selinger, E., & Whyte, K. P. (2012) Nudging cannot solve complex policy problems. *European Journal of Risk Regulation*, *3*(1), 26–31.

Sen, A. K. (1970) *Collective choice and social welfare*. Holden Day.

Shampanier, K., Mazar, N., & Ariely, D. (2007) Zero as a special price: The true value of free products. *Marketing Science*, *26*(6), 742–757.

Simon, H. A. (1955) A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*(1), 99–118.
Starmer, C. (2000) Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, *38*(2), 332–382.
Steffel, M., Williams, E. F., & Pogacar, R. (2016) Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. *Journal of Marketing Research*, *53*(5), 865–880.
Strotz, R. H. (1955) Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*(3), 165–180.
Sugden, R. (2018) *The Community of advantage: A behavioural economist's defence of the market*. Oxford University Press.
Sunstein, C. R. (2015) Nudges, agency, and abstraction: A reply to critics. *Review of Philosophy and Psychology*, *6*(3), 511–529.
Sunstein, C. R. (2016) People prefer system 2 nudges (kind of). *Duke Law Journal*, *66*, 121–168.
Sunstein, C. R. (2017a) Nudges that fail. *Behavioural Public Policy*, *1*(1), 4–25.
Sunstein, C. R. (2017b) People like nudges (mostly). In C. R. Sunstein (Ed.), *Human Agency and Behavioral Economics* (pp. 17–39). Springer.
Sunstein, C. R. (2018) Misconceptions about nudges. *Journal of Behavioral Economics for Policy*, *2*(1), 61–67.
Sunstein, C. R., & Thaler, R. H. (2003) Libertarian paternalism is not an oxymoron. *University of Chicago Law Review*, 1159–1202.
Sunstein, C. R., Reisch, L. A., & Kaiser, M. (2019) Trusting nudges? Lessons from an international survey. *Journal of European Public Policy*, *26*(10), 1417–1443.
Sunstein, C. R., Reisch, L. A., & Rauber, J. (2018) A worldwide consensus on nudging? Not quite, but almost. *Regulation and Governance*, *12*(1), 3–22.
Sürücü, O., Djawadi, B. M., & Recker, S. (2019) The asymmetric dominance effect: Reexamination and extension in risky choice—an experimental study. *Journal of Economic Psychology*, *73*, 102–122.
Thaler, R. H. (1980) Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, *1*(1), 39–60.
Thaler, R. H. (2015) *Misbehaving: The making of behavioral economics*. Norton.
Thaler, R. H., & Benartzi, S. (2004) Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164–S187.
Thaler, R. H., & Sunstein, C. R. (2003) Libertarian paternalism. *American Economic Review*, *93*(2), 175–179.
Thaler, R. H., & Sunstein, C. R. (2009) *Nudge. Improving decisions about health, wealth, and happiness*. Penguin.
Venema, T. A., Kroese, F. M., & de Ridder, D. T. (2018) I'm still standing: A longitudinal study on the effect of a default nudge. *Psychology and Health*, *33*(5), 669–681.
Wakefield, M., Coomber, K., Zacher, M., Durkin, S., Brennan, E., & Scollo, M. (2015) Australian adult smokers' responses to plain packaging with larger graphic health warnings 1 year after implementation: Results from a national cross-sectional tracking survey. *Tobacco Control*, *24*(Suppl 2), ii17–ii25.
Werfel, S. H. (2017) Household behaviour crowds out support for climate change policy when sufficient progress is perceived. *Nature Climate Change*, *7*(7), 512–515.
White, V., Bariola, E., Faulkner, A., Coomber, K., & Wakefield, M. (2015) Graphic health warnings on cigarette packs: How long before the effects on adolescents wear out? *Nicotine and Tobacco Research*, *17*(7), 776–783.
Whitman, D. G., & Rizzo, M. J. (2015) The problematic welfare standards of behavioral paternalism. *Review of Philosophy and Psychology*, *6*(3), 409–425.
Wynes, S., Nicholas, K. A., Zhao, J., & Donner, S. D. (2018) Measuring what works: Quantifying greenhouse gas emission reductions of behavioural interventions to reduce driving, meat consumption, and household energy use. *Environmental Research Letters*, *13*(11), 113002.

How to cite this article: Congiu, L., & Moscati, I. (2022). A review of nudges: Definitions, justifications, effectiveness. *Journal of Economic Surveys*, *36:*188–213. https://doi.org/10.1111/joes.12453